POPULATION MIXING AND THE GEOGRAPHICAL EPIDEMIOLOGY OF CHILDHOOD LEUKAEMIA AND TYPE 1 DIABETES IN NEW ZEALAND

A thesis submitted in partial fulfilment of the requirements for the Degree of

Doctor of Philosophy in Geography at the University of Canterbury

by Laura J. Miller

2008

Table of contents

Abstract

Over the past twenty years the incidence of both childhood acute lymphoblastic leukaemia (ALL) and type 1 diabetes have risen in many developed countries, including New Zealand. Although the explanations for this increase and the precise aetiology of both diseases remain unclear, environmental factors are thought to be important. One factor receiving increasing attention is the role of infections introduced through population mixing. However, previous studies on this topic show mixed results and population mixing itself tends to be under-theorised. Furthermore, this issue has not been adequately assessed in New Zealand, a country characterised by high levels of population mobility.

In this research, a variety of population mixing measures for small areas in New Zealand were developed. National data on ALL registrations were obtained from the New Zealand Cancer Registry, and regional type 1 diabetes data were obtained from the Canterbury Diabetes Register for the Canterbury Region of the South Island. The analyses were undertaken in three stages. First, standardised incidence ratios of each disease were examined at different geographical and temporal scales, between areas of differing socioeconomic status, and in urban and rural New Zealand. Second, cluster analysis was employed to test for spatial-temporal clustering of the two diseases. Finally, multivariate regression analyses were utilised to investigate the association between each disease and the various measures of population mixing at the area-level.

The results reveal similarities in the geographical epidemiology of childhood ALL and type 1 diabetes in New Zealand. The majority of the findings were suggestive of an infectious aetiology for both diseases. In addition, higher incidence of both diseases was observed in areas which increased the most in population mixing over short time periods (6/7 years). Furthermore, raised type 1 diabetes incidence was also associated with high population mixing in early life.

List of tables

List of figures

Acknowledgments

I have thoroughly enjoyed the three years that I have spent carrying out the research for this thesis. This incredible experience owes much to the many people who have helped me along the way.

Firstly, I would like to thank my supervisor Jamie for all of his hard work, advice, and encouragement throughout the course of this thesis. I would also like to thank Ross for his wise counsel when Jamie was overseas and for helpful comments on numerous chapters. I have really enjoyed working with you both.

Thank you also to Jinny, Brian and Russell for the use of their diabetes dataset. Jinny, thank you very much for your help with my diabetes chapters and for the many long and encouraging talks we had. I would also like to acknowledge the help of Craig Wright at the Ministry of Health during the acquisition of various datasets and for his help with geocoding. John Thyne is acknowledged for helping with various GIS-related issues and for never loosing his patience. I wish to thank everyone in the GeoHealth Laboratory, past and present, for their help with an assortment of technical issues and their friendship. The geography department as a whole has made me feel very welcome during my stay.

I would also like to thank my group of friends; Esther, Sarah, Peyman, Jamie, Vicky, Io, Jeff, Charlie, Peter, Nora, Phil, Irfon, Francis and other fellow PhD students for making my time in New Zealand so enjoyable. I am sorry that I faded into obscurity in the last few months! Esther, your friendship appeared out of nowhere and means the world to me. Thank you for making me laugh (a lot!) and also for your proof reading skills. Peyman, thank you for making me laugh in many a fraught moment, and Sarah, thank you so much for your friendship and support. I would also like to thank my cousins, Michelle and Tracey, and friends back in Europe, for being so supportive despite the distance.

Thank you also to my parents for putting up with my long absences, and for all of your help, love and support. As always, I could not have done it without you. Finally, thank you Chris, your love and unwavering patience has helped me through many tense moments. In addition, your superb computing skills proved invaluable on numerous occasions. Thank you for everything and for pretending to understand health geography for me!

Chapter 1: Introduction

1.1 The age of migration

Population movements have been occurring at varying geographical scales since 'time immemorial' (Castles and Miller, 1998, p.1). However, migration started to become a truly global phenomenon in the fifteenth century as a result of European colonial expansion and subsequent industrialisation (Castles, 2000). The dominant flows of people up until the mid-twentieth century were from Europe to parts of the Americas, Asia, Africa and Oceania (Boyle et al., 1998). After a period of reduced migration between 1918 and 1945, international migration expanded further in both volume and scope. Economic growth after the Second World War stimulated large-scale labour migration to Western Europe, North America and Oceania from less developed areas, and ended in 1973 with the oil crisis and associated recession. From the mid 1970s onwards, financial investment shifted away from the old centres and transnational forms of production and distribution reshaped the world economy. Consequently, the older industrial countries witnessed new types of inflows, and new immigration countries emerged in Southern Europe, Africa, the Gulf Oil countries, Asia and Latin America (Castles, 2000). In addition, social and political transformations occurring in the late twentieth century were linked to a number of mass population movements. For example, the end of the cold war in the early 1990s led to increased out-migration from the former Soviet Union and substantial population mobility between the successor states (Zlotnik, 1999). The crumbling of apartheid in South Africa; wars, famine and crises throughout Africa; and the change from dictatorship to unstable democracy in Latin America were also related to large population migrations (Castles and Miller, 1998, Newman and Selm, 2003). Thus, since the mid 1970s there have been fundamental changes in both the nature of migration and the driving forces behind it (King, 1995). This most recent period is widely recognised within the literature as constituting a sharp break with the past in terms of migration patterns (Massey et al., 1998). Castles and Miller (1998) describe the end of the twentieth century and the beginning of the twenty-first century as an 'age of migration' (p.1). People are now moving in larger numbers, at a faster pace, and over greater distances than ever before (Carballo et al., 1998).

1.2 Geography's role in migration research

Academic concern with migration began in earnest in the wake of the industrial revolution (Boyle et al., 1998). This interest in population movements flourished with the advancement of

data sources like national censuses (Pooley and Whyte, 1991). Pivotal research conducted by Ravenstein at the end of the nineteenth century developed eleven principal laws of migration based on empirical observations (Ravenstein, 1885, Ravenstein, 1889). Ten of the eleven laws have been shown to be more or less accurate for the nineteenth century, and many still apply today, attesting to the lasting importance of Ravenstein's work (Boyle et al., 1998). Since this time geographers, economists, sociologists and others have all contributed to what is now a vast body of work attempting to theorise the complex phenomenon that is migration (Pooley and Whyte, 1991).

Early work on migration by geographers focussed upon the measurement and prediction of geographical flows of people, and the motives for migration (Dwyer, 1999, Robinson, 1996). Migration was viewed as an event, and migrants as rational actors responding to economic stimuli. There was thus an emphasis on macro-scale influences and the development of universal laws and hypotheses of migration (Halfacree and Boyle, 1993), like those of Ravenstein (1889) and Zelinksy (1971). While this deterministic approach has been described as a 'logical response' to the bewildering complexity of human movements, it has also received considerable criticism (McHugh, 2000, p.72). In an attempt to uncover broad patterns of, and explanations for population movements, the 'human' in human migration tended to be neglected (McHugh, 2000). Research has commonly ignored that migrations are a cultural experience rich in meaning for individuals, their families and wider communities (Fielding, 1992). Traditional, migration research has also been criticised for being overly empirical and pre-occupied with demographic events, and for uncritically accepting data classifications and results (White and Jackson, 1995). Humanistic accounts of migration were developed in order to address some of these criticisms, however, this resulted in an unhelpful dichotomy between the two approaches (Boyle et al., 1998, White, 1980).

Related to these criticisms, a lively debate developed in the early 1990s calling for more engagement between population geography in general, and migration research in particular, with social theory (Findlay and Graham, 1991, Halfacree and Boyle, 1993, White and Jackson, 1995). Consequently, more integrated accounts of human migration have emerged (Massey et al., 1993). For example, structuration theory has been utilised in an attempt to overcome the determinist/humanist dualism (Halfacree, 1995) and biographical approaches which complement this perspective have increasingly been applied to migration research (e.g. Findlay and Stockdale, 2003, Halfacree and Boyle, 1993). In addition, transnational theories of international migration have been developed which seek to understand and theorise the complex networks of

connection and social fields that migrants forge between distant places (e.g. Bailey, 2001, Bailey et al., 2002, Walton-Roberts, 2004). Feminist research has also contributed to bringing several social theoretical themes to the fore in migration work (Silvey, 2004a, Silvey, 2004b), and recent studies have sought to grapple with the often neglected political, economic, social and cultural consequences of migration (Hugo, 2007, Robinson, 1996). According to Boyle, Halfacree and Robinson (1998), recent migration research is at the forefront of theoretical developments in geography as a whole. It has considered to a greater or lesser extent all of the main philosophical currents debated in geography and the other social sciences.

1.3 The geographies of health

One important consequence of migration is its impact upon population health. However, reviews of the contemporary research agenda in population geography reveal a dearth of studies examining health and migration and their complex interactions (e.g. Hugo, 2007). The majority of geographical work on health and migration has been carried out by health geographers. The sub-disciplines of population geography and health geography share much common ground in terms of their subject matter (population mortality/morbidity and fertility) and the theoretical and methodological issues and debates which they have engaged in over the past 30 years (Rosenberg, 1998). For example, both sub-disciplines have been criticised for, and have subsequently responded to, their conventional lack of attention to social theory (Graham, 2004, Kearns and Moon, 2002). This criticism has been closely associated with arguments critical of the dominance of positivism and the use of quantitative methods in both health geography (e.g. Dyck and Kearns, 1995, Litva and Eyles, 1995) and population geography (e.g. Halfacree and Boyle, 1993, McHugh, 2000). As previously noted for population geography, health (or medical) geographers have also been accused of reducing individuals and their identities to over simplified categories in statistical analyses (e.g. black or white) (Dyck and Kearns, 1995) or to dots on a map (Brown, 1995), resulting in a neglect of human experience and agency.

Within health geography, the contemporary re-theorisation of the sub-discipline is most clearly reflected in the debate that followed Kearns' paper in the *Professional Geographer* in 1993. This paper attempted to encourage a more cultural/humanistic standpoint within health geography and argued for a renewed role for place in a 'post-medical' geography of health. According to Kearns, inadequate attention had previously been given to understanding place as an experienced zone of meaning and familiarity in health research (Kearns, 1993). Since this paper there has been a general broadening in the focus of health geography research from the narrow concerns of

disease ecology and disease services, to wider social models of health, well-being and health care (Kearns and Moon, 2002). Accordingly, this evolution has resulted in a diversification of research topics and methodologies (Parr, 2002), and has reconnected the sub-discipline with theoretical developments occurring in mainstream geography (Brown and Duncan, 2000). For instance, there has been a renewed interest in geographies of the body within health geography (Dorn and Laws, 1994, Dyck, 1995, Parr, 2002) which parallels work in geography as a whole (Longhurst, 1994). An example of this work includes research by Wilton (1999) who used qualitative methods to uncover the embodied experiences of people living with HIV/AIDS in Los Angeles. There has also been an increasing focus upon the role of place (rather than space) and contextual factors on influencing patterns of health and health service delivery (Gesler, 1992, Kearns and Gesler, 1998). For example, analysis of health care delivery has recently drawn upon work on therapeutic landscapes. Conradson (2005), for instance, has demonstrated that the effectiveness of respite centres for the physically impaired can be partly explained by their scenically attractive natural setting and for being identified as places of healing and relaxation. The considerable breadth of topics now researched within health geography is evident in the collections of papers presented at recent international health geography conferences (e.g. Earickson, 2007, Smyth and Thomas, 2005).

While the 'old' focus upon disease ecology has consequently received much criticism, it has also been recognised that there needs to be space for difference. Indeed, many health geographers have emphasised the importance of continuing quantitative health research informed by positivist perspectives (e.g. Bennett, 2005, Kearns and Moon, 2002, Mayer and Meade, 1994). It has been argued that the old and new developments in the field should be interpreted as complimentary rather than competitive, with different approaches each adding to our understandings of health (Curtis and Taket, 1996). Moreover, quantitative methods can engage in important aspects of social theory and have much to offer key research themes in health geography. For example, one means of incorporating social theory within a quantitative framework involves the use of multilevel modelling (Duncan et al., 1996). This approach draws upon structuration theory by recognising the importance of an engagement between both individual human agency and higher-level social structures for a full understanding of health problems. Multilevel modelling therefore offers a potentially 'rich analytical perspective' on the geography of health (Gatrell, 2002, p.69). A growing body of research now exists which employ multilevel models to investigate the relative importance of individual, household and place-level attributes on variations in health outcomes and health-related behaviours (e.g. Duncan et al., 1996, Haynes et al., 2003, Jones and Duncan, 1995, Moon and Barnett, 2003, Shouls et al., 1996). Additional

work which seeks to integrate theory into predominantly quantitative health research methodologies is discussed by Cummins et al (2007), who argue in favour of quantitative health research which incorporates relational views of space and place, and thus draws upon actor network theory. Further examples of recent innovative quantitative work includes, but is not limited to, research into the modelling of complex disease epidemics (e.g. Smallman-Raynor and Cliff, 1999), improved estimation of environmental exposures (e.g. Kingham et al., 2008), and using geographical information systems to measure community resource accessibility for small areas to better understand the contextual influences on health outcomes and behaviours (e.g. Pearce et al., 2008, Pearce et al., 2006b).

Quantitative approaches also have important advantages from a public health perspective. When examining illnesses where the causes are unknown (e.g. childhood leukaemia), quantitative methods can help to reveal new insights into disease aetiologies, by assessing the strengths and directions of associations between diseases and potential individual and wider-level risk factors. In addition, these approaches are essential for estimating the extent of the relationships between risk factors and disease. For example, such methods can quantify the magnitude of the associations between features of specific places and health outcomes. Consequently quantitative approaches can determine how far relationships between health and places are generalisable or variable across whole populations (Cummins et al., 2007). Moreover, the use of geographical quantitative techniques can identify specific areas with high incidence of disease or various health-related behaviours, and can accordingly aid health care planning initiatives. Quantitative frameworks can be employed to help policy makers decide on the type of health care services required, and where to optimally locate them (Curtis and Taket, 1996). For example these approaches have been applied to locating the optimal site for a cervical cancer unit in Trent, England (Smallman-Raynor et al., 1998) and private health care services in India (Kumar, 2004). In geographical research more widely, Dorling and Shaw (2002) have recently stressed the importance of quantifying arguments for policy-related work in order to show how much a particular problem matters and how it can be influenced.

Whilst new perspectives in the geography of health have undoubtedly enriched the sub-discipline, quantitative approaches remain crucial for addressing many health issues of contemporary concern. These approaches are continuously improving and are increasingly drawing upon social theory in order to reveal a more nuanced perspective on health issues and their relation to place. Recent theoretical and methodological debates within the discipline of geography as a whole (and influences from outside the discipline), have broadened the scope of

research conducted within both population and health geography. Consequently, there is now a range and diversity of approaches available to geographers, each able to add to our multiple interpretations of health and migration.

1.4 Migration and health

The study of migration between places and over time is a 'further conceptual dimension that the consideration of place can contribute to the understanding of health' (Tunstall et al., 2004, p.8). It is widely recognised that migration, health and place are intrinsically linked. The movement of people to new environments can have both positive and detrimental impacts on the health of those who move, those they join, and those they leave behind (McKay et al., 2003). Furthermore, health can impact upon people's propensity to migrate and where they migrate to (Bentham, 1988). Analysis of the consequences of migration upon health has traditionally focussed upon the role of population movements in the spread of communicable diseases (Cliff et al., 1986, Gellert, 1993, Longini et al., 1986). In addition, there is a large body of literature on the extent to which individuals moving to another country adopt the health profiles and ill-health risks of those who have always lived there. Numerous studies have shown that over time the health profiles of immigrants become similar to those of the host population, often as a result of adopting similar diets and lifestyles (e.g. McCredie et al., 1999, Moradi et al., 1998, Singh and Miller, 2004, Syme et al., 1975). However, other studies do not support this 'acculturation' hypothesis (e.g. Balarajan and Bulusu, 1990, Harding et al., 1996). More recently, work has been conducted on the mental health issues faced by immigrants in their new environments (e.g. Mirsky et al., 2007, Warfa et al., 2006), especially among refugee populations (e.g. reviewed in Aldous et al., 1999). Furthermore, longitudinal analysis of health selective migration has contributed to the understanding of inequalities in health and mortality between places (e.g. Boyle and Duke-Williams, 2004, Brimblecombe et al., 2000, Norman et al., 2005). For example, health selective migration has been shown to exacerbate geographical inequalities in mortality at the district-level in Britain between 1991 and 1996 (Brimblecombe et al., 2000). Whilst the majority of past research on migration and health has been quantitative, qualitative methods are also progressively being employed (e.g. Dyck and Dossa, 2007, Elliott and Gillie, 1998).

One increasingly influential area of research in the field of migration and health is the role of population mixing in influencing childhood health. Population mixing is part of everyday life and yet its exact definition remains elusive. It involves the movement and interaction of people over varying temporal and spatial scales. Consequently, it can include anything from people

meeting in the local corner shop to the transnational residential relocation of entire families. Through person-to-person contact, population mixing has the potential to introduce infections to susceptible populations. As well as having consequences for the spread of communicable diseases, population mixing has been linked with a number of non-communicable diseases, such as childhood leukaemia, type 1 diabetes, sudden infant death syndrome and multiple sclerosis (Bentham, 1994, Dean, 1971, Greaves, 1997, Kolb and Elliot, 1994).

Research concerning population mixing and childhood health began in the early 1980s in the United Kingdom after apparent clusters of childhood leukaemia were reported around the nuclear reprocessing plants of Sellafield and Dounreay (Kinlen and Doll, 2004). Considerable public concern was raised regarding the initial assumption that radiation from the plants was the cause of the excess leukaemia cases (Gardner, 1991). However, subsequent research showed that these cases could not be explained by levels of radiation in the environment (Black and Independent Advisory Group, 1984, Committee on Medical Aspects of Radiation in the Environment, 1988) or by paternal occupation at the plants (Parker et al., 1993, Urquhart et al., 1991). The epidemiologist, Leo Kinlen, noted that the increases in leukaemia occurred after an influx of migrant professional workers to these previously remote areas (Kinlen, 1988). Consequently, Kinlen hypothesised that childhood leukaemia is a rare response to a common infection introduced through population mixing to non-immune children. Kinlen tested his 'population mixing hypothesis' by examining leukaemia mortality in Glenrothes, a new town in rural Scotland with no nuclear plants and which experienced a doubling of its population in the 1950s and 1960s. A significant excess of leukaemia deaths (7 observed, 1.5 expected) was found in children aged under 5 years for the period of greatest population growth (Kinlen et al., 1990). Following on from this study, he and others have examined unusual population mixing events in the UK (e.g. Kinlen, 2006, Kinlen and Balkwill, 2001, Kinlen et al., 1993) and Europe (Kinlen and Petridou, 1995) with consistently supportive findings.

As well as being increasingly linked with childhood leukaemia, population mixing has also more recently been implicated in the aetiology of childhood type 1 diabetes (Parslow et al., 2001). Childhood leukaemia and type 1 diabetes have many epidemiological and potential aetiological factors in common, including a possible causal role for infection (Feltbower et al., 2004, Greaves, 2007, Staines, 1996). However, of the few studies which have examined the role of population mixing on the incidence of type 1 diabetes, all show increased risk of this disease to be associated with *low* levels of population mixing (Cox, 2007, Parslow et al., 2001). In addition to the mixed nature of recent findings on population mixing and child health, many population

mixing studies can be criticised for using theoretically under-developed measures, and for failing to adequately address the important issue of timing within the study design. Furthermore, there is geographical bias in the studies conducted to date, with the preponderance of work carried out in Europe, and especially the UK. Thus, population mixing also needs to be examined in other geographical settings, especially those with high international in-migration and high internal population mobility. New Zealand is one such setting.

1.5 The New Zealand context

Historically, New Zealand has been referred to as a traditional immigration country (Castles and Miller, 1998). New Zealand was initially settled by people from the Pacific Islands between 600 and 800 years ago (Anderson, 2006a) who, from the late eighteenth century, became known as Māori (Bedford et al., 2002b). Since this time, New Zealand has gained the majority of its population through immigration from Europe (Castles and Miller, 1998). Today, migration flows into the country are increasingly diverse due to a change in immigration policy in 1986 which abolished national origin as a selection factor for immigrant entry (Bedford et al., 2002b, Ongley and Pearson, 1995). Despite its geographically isolated position, New Zealand currently receives over four million international arrivals each year (Statistics New Zealand, 2007c), and around 22.9 percent of people currently living in New Zealand were born overseas (Statistics New Zealand, 2006c). Moreover, New Zealanders are becoming increasingly mobile within the country, with almost 60 percent of the population changing their usual residence at least once between 2001 and 2006 (Statistics New Zealand, 2006b).

A significant volume of research has been conducted in New Zealand on internal and international flows of people, with much of this work carried out by geographers. The majority of research has concentrated on documenting and understanding the main migration streams into and within the country (Bedford et al., 2002b, Friesen, 2003, Heenan, 1979, Heenan, 1985, Newell, 2002), including for instance research on Māori rural to urban population shifts after the Second World War (Barcham, 1998, Metge, 2004). Furthermore, much attention has also been given to Trans-Tasman migration streams (e.g. Bedford et al., 2003, Hugo, 2004). Other research themes include the effects of globalisation on migration flows (Bedford et al., 2002a), issues related to immigration policy (Bedford et al., 2001, Bedford et al., 2005), and migration and multiculturalism (Ward and Masgoret, 2008).

Although this work provides detailed insights into the nature of migration flows into and within New Zealand, it does not consider the possible health implications of such population movements. In terms of the health and migration research conducted in New Zealand, there has been little previous input from geographers. Traditional health and migration research has tended to focus upon the health of Pacific Island migrant groups. For example, there has been a considerable amount of work published on the health of Tokelauan populations who migrated to New Zealand following a hurricane in 1966. These studies collected longitudinal data and thus assessed health changes in these migrants over time. The results revealed that the migrant Tokelauans' blood pressure levels increased after their move to New Zealand and were much higher than in Tokelauans who remained on the atolls (Beaglehole et al., 1979, Prior, 1992, Prior, 1975, Salmond et al., 1985, Ward et al., 1980). Furthermore, rates of type 2 diabetes were significantly higher, and increased more rapidly, in the migrant group compared to the non-migrant group between 1968 and 1982 (Østbye et al., 1989). These studies point to the convergence of health status between migrants and the destination host population after migration. A strength of this work is that it was able to control for possible selective migration and it is thus internationally renowned and still widely referred to today (Elford and Ben-Shlomo, 2004, Gatrell, 2002). More recent migration and health research involving Pacific peoples includes examinations of cancer trends (Tukuitonga et al., 1992) and mental health disorders (Foliaki et al., 2006). An increasing literature is beginning to emerge on the health of Asian migrants in New Zealand (e.g. Abbott et al., 2003, Abbott et al., 1999, Cheung and Spears, 1995, Ho, 2004, Pernice et al., 2000, Rasanathan et al., 2006), reflecting the growing number of Asian people moving to New Zealand since policy changes in 1986 (Bedford et al., 2002b). Migrant health care is also beginning to receive attention in the New Zealand setting (e.g. DeSouza, 2005, Hobbs et al., 2002), although additional work is warranted.

Further research is also necessary on the potential role of population mixing in influencing child health in New Zealand. To date, only one limited study has attempted to investigate this association in New Zealand. This study examined leukaemia rates and population growth in just three small areas in the North Island of the country, and was based on a small number of cases (Dockerty et al., 1996a). Furthermore, no studies in New Zealand have examined the potential effects of population mixing on the incidence of childhood type 1 diabetes. Both childhood leukaemia and type 1 diabetes have been rising in New Zealand in recent years (Campbell-Stokes and Taylor, 2005, Dockerty et al., 1996b, Willis et al., 2002b), but the reasons for the increase, and their precise aetiologies in general, remain unclear. Thus in light of the limited

nature of New Zealand research in this area, and the criticisms of previous population mixing studies in general, a more systematic evaluation of this association is required.

1.6 Aims of the thesis

In this thesis a range of quantitative methods are employed to firstly examine the small area geographical epidemiology of childhood acute lymphoblastic leukaemia (ALL) and type 1 diabetes in New Zealand, and secondly to examine the association between each disease and population mixing. More specifically, the aims of this research are to:

1. Examine the incidence of ALL and type 1 diabetes in children in New Zealand.

   Patterns by age at diagnosis, sex, and ethnicity are explored, as are changes in incidence over time.

   Patterns by area of residence at diagnosis are also examined, including the small area geographical distributions of each disease, and variations in each disease by the area-level deprivation and the urban/rural status of areas. In addition, significant clusters of each disease are identified.

2. Examine whether population mixing is associated with either ALL and/or type 1 diabetes at the small area-level in New Zealand. In detail, this research aims to:

   Develop measures of population mixing to capture a range of different types of population movements at the small area-level in New Zealand, over varying time periods.

   Test a modified version of Kinlen's population mixing hypothesis that a large increase in population mixing over short periods of time will be associated with an increased risk of childhood ALL. This theory will also be extended to childhood type 1 diabetes.

   Test the hygiene hypothesis that low levels of population mixing encountered in early life makes children more susceptible to type 1 diabetes.

1.7 Structure of the thesis

The thesis consists of ten chapters and begins by placing population mixing in the wider context of migration and health research. How population mixing has been defined and measured in previous studies of childhood leukaemia and type 1 diabetes is then examined and critiqued (chapter two).

Chapters three and four review the current knowledge on childhood ALL and type 1 diabetes respectively. These chapters summarise the basic epidemiology of each disease and the main arguments and themes associated with their aetiologies. Of specific interest to this research is the potential causal role of infections, and links to population mixing.

With the wider literature reviewed, chapter five turns towards the data and methods used in this study to address its main aims. The first part of the methods section outlines the epidemiological techniques employed to determine the basic epidemiology and geographical distributions of each disease in New Zealand. The second part of the methods section outlines the approaches used to measure population mixing for small areas across New Zealand, and the statistical analyses employed to test whether population mixing is associated with either childhood ALL or type 1 diabetes.

Chapter six presents the results of a descriptive analysis of the geography of population mixing in New Zealand. It considers how the measures of population mixing vary both spatially and temporally. How population mixing differs by types of area in New Zealand, for example urban versus rural areas, is also investigated.

Chapters seven and eight present the results of the analyses that test the associations between population mixing and first ALL (chapter seven), and then type 1 diabetes (chapter eight). The first section of each chapter details the geographical epidemiology of each disease in the study area.

Chapter nine draws together the key findings of the ALL and type 1 diabetes analyses, with a particular focus on interpreting the population mixing results. Chapter ten concludes the thesis by summarising the main outcomes and implications of this research, before making some suggestions for future research directions.

1.8 Conclusions

This research is situated in, and draws upon, the fields of both population and health geography. As this chapter has shown, these sub-disciplines share many of the same strengths and weaknesses and have equally been criticised for their over-reliance on positivism and associated quantitative methods. However, this chapter has argued that quantitative approaches remain highly valuable, especially within the field of health geography. The broad research area of migration and health was then introduced, and the effect of population mixing on childhood health was identified as an important research gap within the literature more widely, and especially within New Zealand. In order to address some of the criticisms of previous population mixing studies, a selection of quantitative methods are employed in this research. The following chapter expands upon some of the migration and health research themes touched upon here, before examining the topic of population mixing in more detail.

Chapter 2: Population mixing and health

2.1 Migration and health

The links between migration and health have long been recognised (e.g. Welton, 1872). Both are socially and geographically patterned, yet the causal relationships between the two are often ambiguous (Gatrell, 2002). This area of research draws upon concepts and methods from several disciplines, including anthropology, demography, sociology, human geography and public health (Jatrana et al., 2005). Consequently, a myriad of literature now exists attempting to clarify associations between various aspects of migration and population health. One classic conceptualisation of this relationship is provided by Hull (1979) who points out that the causal link between migration and health can occur in either direction. Health may affect a person's propensity to migrate and migration may affect the health of individuals (Jatrana et al., 2005).

Key geographical work on the health selectivity of migrants found that the majority of migrants, especially those moving long distances, tend to be young and relatively healthy. However, amongst older people, those in poor health are more likely to migrate. Older people tend to move shorter distances to avoid environmental health hazards or to be closer to medical care (Bentham, 1988, Norman et al., 2005). Such movements can have consequences for the geography of health inequalities (Bentham, 1988, Boyle, 2004, Boyle and Duke-Williams, 2004, Brimblecombe et al., 2000, Pearce and Dorling, 2006, Rogerson and Han, 2002, van Lenthe et al., 2007).

In terms of migration's effect on health, the movement of people to new environments can have both positive and detrimental impacts on the health of those who move, those they leave behind and those they join (Gushulak and MacPherson, 2006, McKay et al., 2003). The actual migration event itself can be stressful and sometimes hazardous with adverse consequences for people's well being (Evans, 1987, Kalipeni and Oppong, 1998, Mirdal, 1984, Mirdal, 2006). Once migrants have reached their destination, environmental factors such as climate, altitude, air pollution, humidity, temperature, the amount of solar and other irradiation, can all exert either positive or negative consequences on a migrant's health (Jatrana et al., 2005, McKay et al., 2003). Social and cultural change, including personal behaviours and lifestyles, can occur after arrival with varying effects for health (Hull, 1979). For example, migration often necessitates fundamental changes in diet and various sanitation habits (Carballo et al., 1998). Societal-level factors such as ethnic/racial prejudice, social segregation and employment discrimination can all have adverse health effects on migrants (Bauder, 2003, Singh and Miller, 2004). However, better

health care in destination areas may improve the health of migrants (Borman, 2004, Jatrana et al., 2005). Related to such themes, there is a growing literature on the extent to which individuals moving to another country adopt the health profiles and ill-health risks of those who have always lived there (Antecol and Bedard, 2006, Gatrell, 2002, Singh and Miller, 2004).

Migrants can also impact upon the health of destination populations. Human movements can aid the spread of communicable diseases, and have done so for a long time (Frenk and Gomez-Dantes, 2002, Frenk et al., 1997, Gellert, 1993). For example, in the eighteenth century, explorations led by Captain Cook decimated indigenous populations in many of the Pacific Islands through the introduction of syphilis, measles and tuberculosis (Weiss and McMichael, 2004). When Europeans arrived to settle in New Zealand from 1790 onwards, they brought with them influenza to which the native Māori people had no immunity, with devastating effects (Crump et al., 2001). Their numbers had almost halved by the end of the nineteenth century (Pool, 1973). However, what is new, is the speed and ease through which communicable diseases can now be spread. Worldwide travel now produces thousands of potentially infectious contacts daily, and improvements in air travel have made even the longest intercontinental flights briefer than the incubation period of any human infectious disease (Frenk and Gomez-Dantes, 2002). A recent example of the rapidity in which disease can spread via international flights, is the outbreak of severe acute respiratory syndrome in 2002-2003. This disease originated in China and then spread to Hong Kong, Vietnam, Singapore, Taiwan, the Philippines, Canada and Germany, infecting thousands and killing around 800 people (WHO, 2006).

As well as being important in the spread of communicable diseases, movements of people have also been implicated in the increased incidence of some non-communicable diseases (Bentham, 1994, Dean, 1971, Greaves, 1997, Kolb and Elliot, 1994). Of central importance to this thesis, is that influxes of new people to previously remote areas have been implicated in clusters of childhood leukaemia. It has been hypothesised that childhood leukaemia results from exposure to an (as yet unidentified) virus introduced to susceptible children through population mixing (Kinlen, 1988). The remainder of this chapter considers what population mixing is, how it has been related to disease occurrence, and critically reviews previous studies on the topic, before giving a summary of the findings to date.

2.2 Population mixing and health - a detailed review

Despite being an example of how movements of people could directly influence the health of host populations, population mixing is rarely mentioned (with the exception of Gatrell, 2002) in reviews of migration and health (e.g. Hull, 1979, McKay et al., 2003). Indeed, population mixing studies themselves tend to ignore the wider migration and health context in which their work is conducted. The term 'population mixing' has been used in many epidemiological studies and yet its definition remains elusive. It is a complex topic; researchers disagree on what population mixing is, how it should be measured and how various results should be interpreted. This section intends to consider how population mixing has been defined in the literature, some of the limitations of these definitions, and how these definitions could be improved.

2.2.1 Population mixing definitions

Although inherently part of people's everyday lives, as a definable concept, population mixing is far from straightforward. There are many different ways in which populations can 'mix'. One example is the migration of people to new residences either within the same country (internal migration) or to another country (international migration). There are also more regular movements, either on a weekly or daily basis, of people who commute to work. Population mixing also occurs when people go to work locally, go grocery shopping and when people partake in leisure time activities. Children mix by playing with neighbours, attending playgroups and later at schools and universities. To complicate matters further, individuals can belong to many different social subgroups and the transition from one subgroup to another (by ageing or migration for example), as well as the contact rates within and between subgroups, will vary according to many different factors. Such factors include socioeconomic status, political, social and historical context, and individual behaviour, and many of these will in turn be confounded by one another (Fine, 1993). An example may be that ethnic minorities, especially those in lower socioeconomic groups, living in large urban centres may remain relatively segregated from the rest of the host society and thus have a limited range of contacts. These few examples illustrate the many different temporal and spatial scales at which population mixing occurs, and some of the possible confounding factors, that make population mixing a complex process to define and analyse.

The first to mention 'population mixing' in relation to leukaemia incidence was Kinlen in his frequently cited 1988 paper in which he developed his now seminal 'population mixing

hypothesis'. In this paper, Kinlen (1988) sets out a theory to explain the increased incidence of childhood leukaemia in rural areas in the UK, subject to extreme population influxes. However, the term is not explicitly defined in this paper. In this, and his subsequent studies, (e.g. Kinlen et al., 1990, Kinlen et al., 1995, Kinlen and Hudson, 1991, Kinlen et al., 1991, Kinlen and John, 1994, Kinlen et al., 1993), he implies that population mixing occurs in rural areas after a large increase in population, and that it involves an increase in the number of infected and susceptible people who will come into contact with each other.

In the literature more widely, the concept of population mixing has been defined in diverse ways (Law et al., 2003, Parslow et al., 2005). Examples in the early work included viewing population mixing as: population growth (Dockerty et al., 1996a, Kinlen et al., 1990, Kinlen and Petridou, 1995, Langford, 1991); increases in commuting (Kinlen et al., 1991); levels of paternal occupational contact (Kinlen, 1997); proportions of servicemen (Kinlen and Hudson, 1991); proportions of construction workers (Kinlen et al., 1995, Kinlen et al., 1993) and proportions of wartime evacuees present in each area (Kinlen and John, 1994). Beginning with Stiller and Boyle (1996), later work began explicitly incorporating the origin of the newcomers to an area as an important aspect of population mixing measures, although Kinlen mentioned its significance as early as 1990 (Kinlen et al., 1990). For example, Parslow et al. (2001), and Feltbower et al. (2005) defined population mixing as the degree of population migration and the extent to which incoming migrants originate from different areas. A slightly more sophisticated definition was offered by Law et al. (2003). They defined population mixing as a measure of the volume of people who moved, and the diversity of their origins, but they also made it explicit that they were measuring population mixing due to residential mobility (as opposed to other, more temporary movements of people) and that they were using population mixing as a proxy for the mixing of infectious disease.

The majority of definitions are data-focused and lack theoretical explanation and depth. An important facet that any definition of population mixing should include is that of potential person to person contact. As a result, a useful definition of this term may be: the movement and *interaction* of people over time and space. It should also be noted that population mixing occurs differentially according to a number of social, economic and cultural factors such as those outlined above, and also in the wider migration literature (Boyle et al., 1998). Thus it is important to determine which groups of people are most likely to move, which areas in New Zealand experience the most population mixing, and how levels of population mixing change

over time. This topic is important since population mixing has been implemented in the aetiology of a number of diseases.

2.2.2 Population mixing and disease

Disease in humans occurs as a consequence of an interaction between a human host, an infectious or other type of agent, and the environment that promotes the exposure. For such an interaction to take place, the human host must be susceptible. Susceptibility is determined by a combination of factors including a person's genetic background and immunological and nutritional characteristics (Gordis, 2000). The amount of disease in a population therefore depends on a balance between the number of people who are susceptible and therefore at risk of getting the disease, and the number who are not susceptible, or immune to the disease. Of central importance is 'herd immunity', which is where resistance to an infection or disease in the non-immune section of a community occurs as a result of a high enough proportion of immune members in that community. This so-called 'herd effect' (John and Samuel, 2000) occurs because once a certain proportion of a community is immune, the likelihood that an infected person will come into contact with a susceptible person, is much smaller. Many more of the infected person's contacts will be with immune people, to whom the disease cannot be transmitted (Gordis, 2000).

Effective contact between infected and susceptible individuals is thus the basis for the spread of any infection (Kinlen et al., 1990). Characteristics of the places where people live can influence the likelihood of such contact (Fox et al., 1971). For example, population density is an important factor in determining the successful spread of an infectious agent (Anderson and May, 1979, Anderson and May, 1982, Wilson et al., 1983). Where densities are high, each primary case of infection tends to generate many secondary cases as a result of frequent contact between infected and susceptible individuals (Kinlen et al., 1990). The size of the dose of the infectious agent and the age at exposure are also important factors in the spread of disease (Anderson, 1982). Kinlen, et al. (1990) postulate that leukaemia may result from a large dose of the relevant agent(s) which occur through repeated contacts with infected individuals, and which are more likely when many people contract an infection around the same time.

As previously noted, there are many different ways in which movements of people between areas can promote contacts, which may be relevant for the transmission of disease. Population mixing has the potential to increase the range and level of infections in a community through contacts

between individuals moving into, and within, the area (Parslow et al., 2001, Rhodes and Anderson, 1996). Areas which experience high levels of in-migration and population mixing, such as central urban areas, have been shown to have a greater prevalence and variety of infections endemic in the area, due to a higher proportion of infected and susceptible people (Anderson and May, 1982, Kinlen et al., 1990). Regions subject to lower population densities and migration rates, such as remote rural areas, are less likely to sustain an extensive range of endemic infections, resulting in lower population or herd immunities (Rhodes and Anderson, 1996). Migrants also increase the pool of susceptible people so can encourage the spread of an infection endemic in the host population (Bentham, 1994). Where migrants move further distances, they are less likely to have had previous contact with the local population and they may bring with them infections to which local populations have little or no immunity. In addition, the greater the diversity of origins among sizeable groups of residents, the greater the differences in herd immunity there will be among subgroups (Kinlen et al., 1990). These ideas form the basis of the population mixing hypothesis touched upon previously. Population mixing was first linked to childhood leukaemia incidence in the late 1980s, but has also more recently been associated with childhood type 1 diabetes.

2.2.3 Review of previous studies

The majority of studies carried out to date investigating the associations between population mixing and health focus on childhood leukaemia (Tables 2.1 and 2.2). Most of these studies concentrate primarily on the aetiology of the disease and investigate whether there is evidence for an infectious cause (e.g. Alexander et al., 1997). In many of these studies, little attention has been given to the theory behind population mixing, the methods used to measure it, and the potential consequences of these issues on the results obtained.

2.2.3.1 Area-level measures of population mixing

Ecological studies use proxy measures of exposure to community infections, in order to explain geographical variations in incidence of disease (Feltbower et al., 2005). Ecological studies in general have a number of advantages and limitations all of which need to be considered when interpreting analyses conducted at this level. Firstly in terms of the limitations, the data used are often based on averages for relatively large geographical areas in which much variation can exist. Also, unrecognised confounders can introduce bias into the results. Furthermore, ecological studies are affected by the Modifiable Areal Unit Problem (MAUP) whereby using

different spatial units of analysis can yield substantially different results (Flowerdew et al., 2008). Importantly for the spread of disease, the estimated fixed effects from exposure to infections at the community-level may not directly resemble those at the individual-level (Staines, 2001). However, it has been argued that because of the relevance of herd immunity to the spread of infections, population dynamics need to be assessed at the community-level (Kinlen, 1988, Kinlen et al., 1990, Kinlen et al., 1991). According to Staines (2001) it is difficult to see how population mixing could be measured at the individual-level. Another more practical advantage is that the small numbers of cases that are often witnessed for rare childhood diseases, like leukaemia, often require aggregation to larger geographical units in order to improve the statistical power of the analyses.

Population change

The majority of ecological studies conducted to date have used population growth as a measure of the population mixing occurring within areas. The first study to use population growth as a proxy for population mixing at the community-level was Kinlen's 1988 study of leukaemia mortality in Glenrothes, Scotland. Glenrothes was chosen as it was a new town in a rural area of Scotland with no nuclear plants, which had experienced a doubling of its population in the 1950s and 1960s, and was thus appropriate for testing the population mixing hypothesis. A significant excess of leukaemia deaths (10 observed, 3.6 expected) was found in those aged below 25 years in Glenrothes for the period 1951 to 1967, the period of greatest growth. The raised number of leukaemia deaths was mainly due to an excess in deaths in the under 5 years category (7 observed, 1.5 expected). Between 1968 and 1985 fewer deaths than expected (5.18) were recorded (1) and other areas showed no significant excess in leukaemia deaths for any period. Following on from this study, Kinlen et al. (1990) examined other new towns in the UK. These were separated into two categories 'overspill' (the majority) and 'non-overspill' (rural new towns with incomers from diverse origins). As in Glenrothes, a significant increase in deaths from leukaemia was observed (observed to expected ratio of 2.75) in those aged under five years in the non-overspill rural new towns in the first half of the 40 year period, but there was no excess in the second half. An increase in population density and the diversity of contacts in rural new towns in the 1950s may have produced an epidemic of an unidentified infection which in turn led to an increase in the number of leukaemia cases recorded. However, this increased exposure may also have led to immunising doses being received by a large proportion of the population explaining the deficits of leukaemia in older age groups later on (Kinlen et al., 1990).

As well as being subject to the more general limitations of ecological study designs, the measure of population mixing used in these studies is far from ideal. Population change is not a direct measure of migration but simply a measure of the net loss or gain of a population over a certain period. Thus areas which experienced a net loss of population (due to deaths and/or out-migration) may also have experienced a sizeable inflow of new migrants. As long as the necessary mix of infected and susceptible people is present in such areas, increases in the incidence of childhood leukaemia may still be witnessed. Being retrospective studies, areas were identified which the researchers thought would be subject to excess or fewer cases of leukaemia compared to the national average due to a prior hypothesis, rather than testing the association across the whole country. Other areas with a high incidence of childhood leukaemia were probably missed. In addition, these studies assumed a threshold effect; they only considered areas in which *considerable* growth had occurred. However, later studies (Kinlen and Bramald, 2001, Kinlen et al., 2002, Stiller and Boyle, 1996), found excesses of leukaemia in situations falling well short of the extreme population mixing events examined in these early studies.

Langford (1991) addresses some of these issues by considering population change from 1961 to 1971 in all of the local authority areas in England and Wales. In areas where the population increased by more than 50 percent, significant increases in leukaemia were found in children aged 0-14 years (relative risk of 1.41), relative to all other categories of population change. These excesses were most marked in rural areas concentrated around the major conurbations of London, the Midlands and Merseyside (although this was not statistically tested). This result tends to contradict Kinlen et al.'s (1990) finding that overspill areas have fewer leukaemia cases due to mixing of similar populations. A more recent study in the US also examined population change in counties in Iowa, New Mexico and Utah (Wartenberg et al., 2004). For the period 1980 to 1989, the relative risk of acute lymphoblastic leukaemia (ALL) increased as a function of population change from 1.9 in US counties that had increased in population by up to 10 percent, to 2.6 in those which had increased by greater than 20 percent. However, all of the sub-analyses were subject to small number problems and the geographic resolution of this study is likely to conceal large variations in population growth rates.

More recent work by Alexander et al. (1997) used population growth within smaller geographical units (census areas) as a proxy for population mixing in their study of leukaemia clusters in Hong Kong. Each area was ranked by population growth and the ranked groups were then divided into ten categories with approximately equal total child-years at risk. The areas in the tenth category were selected as areas of extreme population increases. Statistically significant

clustering of ALL (p-value <0.007) was found in the areas which witnessed the most extreme population mixing between 1981 and 1986. While this study is useful in that it was the first to consider the issue of population mixing and leukaemia in Asia, the use of population growth as a proxy for population mixing is inadequate. Furthermore, this study is subject to all of the drawbacks inherent with ecological studies, and it fails to control for other possible confounding variables like population density or possible radiation exposure.

Only one study regarding the relationship between population mixing and childhood leukaemia has been conducted in New Zealand (Dockerty et al., 1996a). This research also used population growth as a proxy for population mixing. Age-adjusted rate ratios for leukaemia showed no significant relationship with periods of population increase in three areas of the North Island when compared to the rest of New Zealand. However, this research is subject to a number of limitations. Similar to Kinlen's studies in the UK (Kinlen, 1988, Kinlen et al., 1990), this work only concentrated on three study areas in the North Island of New Zealand (Rotorua, Matamata, and Whakatane) making it limited geographically. It thus left other areas in New Zealand with potentially significant excesses of leukaemia unidentified, by assuming a threshold effect existed. It also utilised an inadequate measure of population mixing as it simply assessed large population increases in these areas. These increases could be partly due to an increase in births rather than just in-migration. Furthermore, any new migrants entering these areas may have originated from nearby towns, and thus would not bring in new infections which locals may have lacked immunity for. This study also had limited statistical power (Kinlen, 2000) due to small numbers of cases and did not control for potential confounders, such as population density which is critical to whether an infection can maintain itself. Finally, it was a poor test of Kinlen's hypothesis since one of the study areas, Rotorua, would have experienced population mixing through tourism prior to the study period.

Employment-related moves

Studies have also used data on employment-related movements as a measure of population mixing. Such studies tend to involve considerable numbers of people that are relatively temporary in nature. These studies define population mixing as the proportion of new workers in each area in relation to the local population. The numbers of new workers represent a percentage of the new in-migrants into each area and thus some of the problems inherent with using net population growth as discussed above, are overcome.

Kinlen et al. (1993) calculated the proportion of new oil workers (from oil company records) per 100,000 economically active men in each postcode sector in Scotland during a period of rapid economic growth in the North Sea oil industry. They found that in rural areas with the greatest proportion of oil workers, there was a significant excess of leukaemia (31 observed, 16.6 expected) in those aged 0-4 years directly after the large workforce increases. These excesses were highest in areas of relatively high social class. Since the oil workers were under stringent conditions of employment which restricted them to the worksite area, there was only an unusual degree of mixing between them and the local men, pointing to transmission of infection by adults back to their home communities. The Dounreay-Thurso area also had high proportions of oil workers and the well known cluster of leukaemia cases occurred just after the surge in the oil industry in the area, suggesting that population mixing, and not radiation from the nuclear plant, had a part to play in the increased leukaemia incidence (Kinlen et al., 1993). Whilst the measure of population mixing used in this research is an improvement on previous studies, it does not account for where the migrant workers originated. The authors note that many of the workers lived outside Scotland especially in the Tyneside and Teesside areas of England, but they do not incorporate these origins into their analyses.

A study conducted in France (Boutou et al., 2002), investigated the association between population mixing and leukaemia around the La Hague nuclear waste reprocessing plant after a period of high influx of construction workers. A population mixing index was defined as the number of male construction workers born outside the French department of La Manche who were recorded as living in the commune, divided by the number of men aged 20-59 years at the 1975 census. The communes were then stratified by their urban/rural status. A positive trend in leukaemia incidence was observed among the rural strata with increasing population mixing. This study thus only considered construction workers who were originally from outside the department/region, and did not include local workers. However, as with Kinlen et al.'s (1993) population mixing measure, it did not account for other migrants who may have entered the area for purposes other than construction work. The possible effects of radiation were also not considered (Boutou et al., 2002).

National military service provides another opportunity for population mixing, and was examined by Kinlen and Hudson (1991) using data from the 1951 census. Proportions of servicemen in relation to all men residing in each of the local authority districts in England and Wales in 1951 were computed. The ratios of observed to expected deaths from leukaemia were expressed as ratios relative to the value for the group with the lowest proportion of servicemen. In rural areas

with the highest proportion of servicemen, a significant excess (observed to expected ratio 1.65) of leukaemia in children under 14 years was observed for the period 1950-53. This excess was highest in children under 1 year (observed to expected ratio 2.14) suggesting an intrauterine infection. The findings point to infection transmitted among adults probably promoted by the cramped conditions of military camp life (Kinlen and Balkwill, 2001, Kinlen and Hudson, 1991).

All of these examples of employment-related population mixing measured relatively large movements of people, usually around the same time, and often leaving the area once the work was completed. Studies which test the association between population mixing and more regular, day-to-day movements of people for employment have also been conducted. The first was carried out by Kinlen et al. (1991) to test the theory that increases in commuting, as a proxy for population mixing increases, promote contacts between susceptible and infected individuals, and are thus related to excesses of leukaemia. Data were extracted for county boroughs in England and Wales from the 1971 and 1981 censuses on the number of residents who worked outside the borough, and the numbers of residents from other areas who worked within the boroughs. The authors then calculated the increase in commuting by area, between 1971 and 1981. A significant excess of leukaemia (observed to expected ratio 1.5) at ages 0-14 was present only in the decile with the greatest increase in commuting. However, the authors did not control for possible confounding by population density and as they note, the effect witnessed may have been compounded by associated population increases.

A more recent study (Stiller and Boyle, 1996), also examined the relationship between commuting and leukaemia incidence. In this research, possible relations between leukaemia incidence rates and commuting (and other population mixing and socioeconomic variables, see below) were investigated using Poisson regression models. Commuting was calculated as the proportion of the total economically active population in each district in England and Wales who were travelling to work in another district, and the proportion who were travelling to work in another district by public transport in 1981. The analyses revealed no significant association between commuting levels and leukaemia incidence. However, this study differed to Kinlen et al.'s (1991) earlier work as it measured proportions of commuters present at the time, rather than increases in commuting over a given period. As the authors note, the districts with high commuting levels as of 1981, were likely to be those where numbers commuting to work had been high for a long time (for example, areas of south east England).

Residential mobility

Employment-related movements of people are often short term. A longer term measure of population mixing that has been used in studies, is the proportion of people moving to live in a new area. There are two main area-based measures of population mixing which have been employed in relation to residential migration. The first is the volume of in-migration into an area (also known as the volume of population mixing), and the second, more complex measure, is an index of the diversity of origins of these in-migrants (or, diversity of population mixing).

The volume of in-migration as calculated by various studies (Dickinson et al., 2002, Dickinson and Parker, 1999, Law et al., 2003, Stiller and Boyle, 1996), measures the total number of people moving into an area, in a given period of time, as a proportion of the area's total population. For example, in Law et al.'s (2003) recent study of population mixing and childhood cancer in Scotland, England and Wales, the volume of population mixing was calculated as the proportion of the population with a different address one year before the census, excluding those who moved within the same electoral ward/postcode sector. The advantage of this measure is that unlike population growth, it is not affected by the births and deaths that occur in the region. It measures actual flows of people into the area as a proportion of the people who are already there. The volume of in-migrants captures movements that are potentially more permanent in nature than those which are purely employment-related. Studies which have utilised this technique, have also created separate measures for child migration (typically under 15 years of age), adult migration (15 years plus) and any-age migration (all ages) (e.g. Law et al., 2003, Parslow et al., 2001, Stiller and Boyle, 1996). This distinction is useful to assess the importance of mixing between various age groups in association with disease incidence rates.

However, simply assessing the number of in-migrants that enter an area has been deemed ineffective (Stiller and Boyle, 1996) since it does not take into account where the migrants originated. If, for example, all of the newcomers into an area are from adjacent areas, it is possible that these populations will have been in contact with each other previously, carry similar infections and/or have similar levels of immunity. Fortunately, this measure also has the potential to examine incidence of ill-health in relation to the proportion of migrants in an area by the distance that they have travelled. For instance, Dickinson et al. (2002) investigated the association between population mixing and childhood leukaemia and non-Hodgkin's lymphoma (NHL) in census wards, in England and Wales for the period 1966-1987. In this study, measures of migration at six levels were calculated for each ward. The proportion of newcomers from

outside: Great Britain; the region; the county; the district; the ward and the proportion of all migrants in the ward, including those moving within the ward, were employed as measures of population mixing. This study found a significantly increased risk of leukaemia and NHL (relative risk 1.9) with increasing community population mixing at the ward-level, which was more marked for movers from a greater distance. Of course, the boundaries used to define how far the migrants had travelled were relatively arbitrary, and analyses conducted using different geographical delineations could potentially obtain different results.

Distance travelled, however, may not be an effective marker of the range of infections circulating in any given area (Parslow et al., 2001). Kinlen et al. (1990) noted early on that the greater the diversity in geographic origins among residents, the more likely that there will be appreciable differences in herd immunity among different subgroups. Early ways of measuring this effect included grouping areas or towns into categories depending on whether the incomers were mainly from surrounding areas or from a greater range of areas (Kinlen et al., 1990). Stiller and Boyle (1996) progressed this idea further with their use of the Shannon Index of Diversity. This index was developed by ecologists to describe the diversity of different habitats in relation to plant and animal species (Magurran, 1988). The index has now been used to investigate the diversity of origin of incoming migrants and its association with childhood diabetes (Cox, 2007, Feltbower et al., 2005, Parslow et al., 2001) and childhood leukaemia (Dickinson et al., 2002, Dickinson and Parker, 1999, Feltbower et al., 2005, Law et al., 2003, Parslow et al., 2002, Stiller and Boyle, 1996). The index produces a measure of diversity on the basis of the number of origins, and the proportion of individuals coming from those origins, and has been considered as a good representation of population mixing (Parslow et al., 2001). It measures the extent to which some total population (in this case, the total incomers to a particular area) is distributed among its component parts (the origin areas).

In Stiller and Boyle's (1996) analysis of ALL in county districts in England and Wales using this index, the combination of higher migration with greater diversity of origins or distance moved was associated with an increase in leukaemia incidence. The raised incidence of leukaemia at age 5-9 in districts with high proportions of child migrants was consistent with the transfer of viruses at school (Stiller and Boyle, 1996). This study was the first to show that even relatively low levels of population mixing may be important in causing leukaemia which contradicts the assumptions made by Kinlen (1988) who believed that increased leukaemia rates would only be found in areas of *extreme* population mixing.

Following Stiller and Boyle's (1996) work, Dickinson and Parker (1999) and later Dickinson et al. (2002) investigated the effects of the diversity of population mixing and childhood leukaemia, the former in the district of Cumbria in north-west England, and the latter across all census wards in England and Wales. Although both studies found significant positive associations between leukaemias in wards with a high volume of in-migrants from outside the region, they found no significant association with the diversity of their origins. However, in their study of childhood cancer and population mixing in England, Scotland and Wales, Law et al. (2003), found elevated risks of ALL (odds ratio of 1.40) to be associated with areas of the *lowest* diversity of origin of all-age migrants, and found no relationship between ALL and the volume of population mixing. Similar results were recorded by Parslow et al. (2002) which concentrated on the smaller geographic region of Yorkshire in England. Moreover, in one of the few studies conducted to date regarding the association between population mixing and childhood type 1 diabetes (Parslow et al., 2001), a negative trend was witnessed, with the highest rates of childhood diabetes found where the diversity of origin of incoming children was low. A similar trend was noted in a study which considered both ALL and type 1 diabetes in the same analysis (Feltbower et al., 2005).

Results to date using a diversity index to assess migrant origins in relation to disease incidence are thus more mixed than the results of studies which utilised more simplistic measures of population mixing. This finding could be partly explained by differences in the study designs. For example, the lengths of the disease diagnosis periods in these studies ranged from seven years (Law et al., 2003, Stiller and Boyle, 1996), to 22 years (Dickinson et al., 2002). Perhaps more importantly, migrant diversity levels were measured at different points within the diagnosis periods. For instance, Law et al. (2003) analysed ALL cases during the period 1991-1996 and measured migrant diversity in 1991: the beginning of the study period. In contrast, Dickinson et al. (2002) considered childhood leukaemia cases between 1966 and 1987, and measured migrant diversity in 1981, towards the end of the diagnosis period. This example highlights the lack of importance given to the *timing* of population mixing measurements in many of the more recent studies, which is often dictated by data availability.

While the migrant diversity index and in-migration rate measure different dimensions of population mixing, it has recently been argued that different combinations of the two measures may have different implications for the spread of infections (Cox, 2007). Consequently, a new categorical population mixing measure has been proposed by classifying areas as having either above or below average in-migration *and* either above or below average migrant diversity.

Therefore areas can be divided into one of four categories: category one areas have high in-migration from a diverse set of origins; category two areas have high in-migration from a narrow set of origins; category three areas have low in-migration from a diverse set of origins; and category four areas have low in-migration from a narrow set of origins. This variable was employed in an examination of childhood type 1 diabetes in Tayside, Scotland to test the hygiene hypothesis that children living in areas with low population mixing (low in-migration and low migrant diversity) are at increased risk of developing the disease. The results supported this theory, with significantly higher incidence of type 1 diabetes noted in areas which had low child in-migration and low child migrant diversity (Cox, 2007). A positive feature of this measure is that it is useful for incorporating more detail regarding individual in-migrants, into statistical analyses. However, one aspect that all of the studies reviewed in this section are missing is a way of measuring a child's individual exposure to infections brought about by such movements of people.

2.2.3.2 Individual-level measures of population mixing

It is relatively difficult to conduct case-control studies to investigate the effect of population mixing on disease due to a lack of individual-level data. Indeed some might argue against a purely individual-level approach because of the relevance of herd immunity to the spread of infections. The most common means of examining population mixing at the individual-level involves measures of paternal occupational contact.

Following the work already conducted on infectious viruses such as poliomyelitis and cytomegalovirus and parental occupational contacts, Kinlen (1997) tested whether children whose fathers had contact with many different people at work, had a higher incidence of this disease than those whose fathers had fewer occupational contacts. Kinlen used cases of children with leukaemia in areas subject to high population mixing, as previously identified in five of his studies (Kinlen et al., 1990, Kinlen et al., 1995, Kinlen and Hudson, 1991, Kinlen et al., 1991, Kinlen et al., 1993). Paternal occupational particulars were abstracted and then categorised by contact level (low, medium, high or very high). The incidence of leukaemia in children with fathers in very high contact occupations was twice as high as those in low or medium contact jobs, and excesses were found to be the most marked in transport and construction-related occupations. However, no increases were found in a somewhat limited examination of this association outside areas of high population mixing in the general population of England and Wales (Kinlen, 1997). Kinlen and Bramald (2001) offer an improvement on the methods used in

this work, in a national case-control study conducted in Scotland. Birth certificates detailing paternal occupation were traced for cases, and three controls of the same age and sex were randomly chosen from the birth registers of the same county. For those aged 0-4 years, the risk of leukaemia and NHL, adjusted for social class, showed a significant positive trend with increasing paternal contact level in rural (p-value = 0.02) but not urban (p-value = 0.26) areas. A further national case-control study was conducted in Sweden (Kinlen et al., 2002) using similar methods to that of the Scottish study, and recorded similar results, as did another UK based study by different researchers (Pearce et al., 2004).

Fear et al. (1999) investigated the hypothesis that increased exposure to infections, through parental jobs involving high levels of social mixing, would reduce the risk of childhood type 1 diabetes. No overall association between parental occupational social mixing and type 1 diabetes in their children was detected for all age groups. However, a protective effect of maternal high occupational social mixing on children diagnosed below five years of age was found (Fear et al., 1999). This study highlights the importance of maternal occupational contact in the spread of disease, contrary to the focus of other studies carried out to date. Drawbacks with such studies in general include the crude categorisation of the various occupations into low, medium, high and very high contact groups; the frequent neglect of maternal occupation; and the complete neglect of the level of social and recreational contacts of both parents. For example, a father whose usual occupation is classed as having a low range of contacts with other people may also travel around the country entering competitions for various sporting events involving large numbers of people.

2.2.3.3 Individual and area-level measures of population mixing

A study by Dickinson and Parker (1999) in Cumbria in north-west England recognised the importance of both individual and area-level population mixing measures. This study used the place of birth of the parents of each child involved to create an individual-level measure of population mixing. The place of birth of both parents was obtained from their birth certificates and each child was subsequently categorised as having both, one, or neither parents, born outside Cumbria. Community-level measures of population mixing utilised included the proportion of parents born outside Cumbria, the proportion of children who had moved in the year before a census and the diversity of counties of origin of all parents (Shannon Index of Diversity). At the community-level, the results showed a significant 11.7 fold increase in ALL and NHL in children born in the region (excluding Seascale) with the highest level of population mixing (as measured by the proportion of parents born outside Cumbria). At the individual-level, the risk

was higher among children of incomers (i.e. both parents born outside Cumbria) than among children of local residents. Although this study suggests that both individual- and area-level measures of population mixing are important risk factors for ALL and NHL, it did not examine these influences simultaneously. For example, the risk of developing leukaemia may be higher in those children whose parents were born outside Cumbria *and* who were themselves born in areas with high levels of population mixing. Such effects are masked by individual-level analyses.

Table 2.1: Studies on population mixing and childhood leukaemia

Study	**Population**	**Methods & Population Mixing Measure**	**Results**
Kinlen (1988)	0-24 yrs Glenrothes Scotland 1951-1967	Ecological study. Observed and expected deaths from leukaemia were statistically compared in the rural new town of Glenrothes. Population mixing: population growth in a previously isolated rural area	Significant excess of leukaemia deaths was found in this rural new town between 1951 and 1967 which was the period of greatest growth in the population.
Kinlen et al (1990)	0-24 yrs New towns Britain 1946-1965	Ecological study. Observed and expected deaths from leukaemia were statistically compared in 5 rural new towns and 9 overspill new towns. Population mixing: population growth near to urban areas vs previously isolated rural areas	Significant excess of leukaemia deaths found in those aged 0-4 in rural new towns whose residents came from a diverse number of origins.
Kinlen & Hudson (1991)	0-14 yrs Counties Britain 1949-1953	Ecological study. Ratios of observed to expected deaths (relative to the group with the lowest proportion of servicemen) by proportion of servicemen in counties and local authority districts were compared. Population mixing: proportion of servicemen by urban/rural area	Significant excess of leukaemia deaths in children aged 0-14 was found in the rural quintile with the highest proportion of servicemen, 1950-1953. Greatest excess of deaths were in those aged below 1 year.
Kinlen et al (1991)	0-14 yrs 28 towns Britain 1971-1981	Ecological study. Ratios of observed to expected number of cases were compared by % change in commuting. Population mixing: % increase in commuting	Significant excess of leukaemia cases were found only in towns in the decile with the greatest increase in commuting.
Langford (1991)	0-14 yrs England & Wales 1969-1973	Ecological study. Poisson p-values were calculated for each area compared to the national average and aggregated into categories by level of population change. Population mixing: population growth	Significantly increased risk of leukaemia mortality in areas experiencing more than 50% increase in population from 1961-1971
Kinlen et al (1993)	0-24 yrs Scotland 1974-1988	Ecological study. Relative risks calculated for rural and urban postcode sectors in Scotland where low, medium and high proportions of oil workers were present, for 3 time periods. Population mixing: proportion of oil workers by urban/rural area	Significant excess of leukaemia cases in those aged 0-4 in rural areas of Scotland with the greatest proportions of oil workers, only for the early post mixing period; explains the Dounreay-Thurso cluster.
Kinlen & John (1994)	0-14 yrs Rural districts England & Wales 1945-1949	Ecological study. Observed and expected deaths from leukaemia were statistically compared in rural and urban districts in England and Wales, by the ratio of evacuees to local children. Population mixing: evacuee index	Significant excess of leukaemia deaths in rural districts of England & Wales receiving high proportion of war-time children evacuees from London.
Kinlen et al (1995)	0-14 yrs England Scotland & Wales 1945-1995	Ecological study. Observed and expected deaths/cases were statistically compared in areas where large construction projects had been undertaken in rural areas. Population mixing: areas with a peak incoming workforce of 1000 people.	Significantly increased incidence of leukaemia near large rural industrial sites or Scottish hydroelectric counties, while construction was under way.

Study	Population	Methods & Population Mixing Measure	Results
Kinlen & Petridou (1995)	0-14 yrs Europe 1950-1987	Ecological study. Standardised mortality rates compared across countries in Europe. Population mixing: countries with large rural population movements	Greece and Italy were found to have unusually high mortality rates from leukaemia. These were the most rural countries which experienced the highest level of rural movements.
Dockerty et al (1996a)	0-14 yrs Rotorua Matamata Whakatane New Zealand 1949-1983	Ecological study. Age-adjusted rate ratios were calculated for three areas in the North Island in New Zealand which experienced population growth in the 1950s, and compared to the rates for the rest of New Zealand. Population mixing: population growth in rural areas.	Age-adjusted rate ratios of leukaemia were not significantly raised during or after periods of greatest influx of people in the 3 rural areas examined.
Stiller & Boyle (1996)	0-14 yrs England & Wales 1979-1985	Ecological study. Poisson regression analysis was utilised to examine variations in leukaemia incidence at the county district-level, according to measures of migration and commuting, whilst controlling for measures of socioeconomic status. Population mixing: in-migrants and diversity of origin of in-migrants (Shannon Index of Diversity) for child and any age, also commuting to another district & proportion of armed forces by district	Trends for higher incidence of leukaemia at ages 0–4 and 5–9 with the proportion of child in-migrants were noted ($P<0.05$). The combination of higher migration with greater diversity of origins or distance moved was associated with higher incidence in both age groups. No significant trends were found with commuting.
Alexander et al (1997)	0-4 yrs Hong Kong 1984-1990	Ecological study. Standardised morbidity ratios for leukaemia were analysed in small areas and checked for spatial clustering and association with population mixing. Population mixing: population growth deciles	For small census areas with extreme population growth, overall incidence of ALL was raised and there was significant evidence of spatial clustering ($P>0.05$)
Kinlen (1997)	0-14 yrs UK 1950-1995	Case-control study of children with leukaemia identified in 5 previous studies. Occupations of the fathers of those in the highest exposure areas were abstracted and then categorised by level of contact. Population mixing: high contact paternal occupations	The incidence of leukaemia in children with fathers in very high contact occupations was twice as high as those in low-medium contact jobs. Excesses were found to be the most marked in transport and construction-related occupations.
Dickinson and Parker (1999)	0-14 yrs Cumbria England 1969-1989	Leukaemia incidence was analysed using Poisson regression in relation to both individual and community risk factors. Population mixing: Individual-level: place of birth of parents (both, one, or neither born outside the district) Community-level: proportion of parents born outside Cumbria, proportion of children who had moved in the year before a census, & diversity of counties of origin of parents.	Incidence was significantly higher among children born in areas with the highest levels of population mixing (RR = 11.7, CI =3.2-43.0) and was highest among children of the incomers. Thus both individual and community influences were important risk factors. No significant association was observed with the diversity of origin of the incoming migrants.
Kinlen & Balkwill (2001)	0-14 yrs Orkney & Shetland Scotland 1941-1955	Cohort study. Observed and expected deaths were statistically compared in Orkney and Shetland in the war and post-war periods. Population mixing: proportion of servicemen to local people	A significant 3.6 fold increase in leukaemia deaths was observed in the war time period, when local people were outnumbered by troops stationed there.

Study	Population	Methods & Population Mixing Measure	Results
Kinlen & Bramald (2001)	0-14 yrs Scotland 1950-1980	Case-control study. Occupations of fathers were categorised by level of contact, and urban and rural differences, and periods of high and low population mixing were examined. Population mixing: high contact paternal occupations by urban/rural area	Increased incidence of leukaemia was found in children aged 0-4 years in rural areas with increasing paternal occupational contact levels, after adjusting for social class. No effect was found in urban areas.
Koushik et al (2001)	0-14 yrs Ontario Canada 1978-1992	Ecological study. Leukaemia incidence rates were analysed using Poisson regression in relation to population mixing and geographic isolation. Population mixing: population growth by urban/rural status	Incidence of leukaemia increased in rural areas after population growth, especially for ALL and children aged 0-4 years (RR = 1.8). No elevated risk was observed in urban areas.
Boutou et al (2002)	1-24 yrs Nord Cotentin France 1979-1998	Ecological study. Incidence rate ratios were analysed using Poisson regression in relation to influxes of workers for the construction of a nuclear power plant. Population mixing: number of male construction workers born outside the area, by urban/rural status	Rural communes with the highest tertile of population mixing witnessed increased risk of leukaemia. Findings were strongest for 1-6 yrs & ALL (IRR = 5.5).
Dickinson et al (2002)	0-14yrs England & Wales 1966-1987	Ecological study. Poisson regression analysis was utilised to examine variations in leukaemia incidence at the census ward-level, according to measures of population mixing. Population mixing: proportion of incomers from outside; the UK, the region, the county, the county district, the ward and the proportion of all migrants in the ward. The Shannon Index of Diversity was also calculated.	For urban areas only, increased risk with higher levels of inward migration, particularly from outside the region was observed RR = 1.9 (1.2-2.9). Although a marked but non-significant effect was seen in affluent rural areas. No significant association was observed with the diversity of origin of the incoming migrants.
Kinlen et al (2002)	0-14 yrs Rural Sweden 1958-1998	Case-control study. Occupations of fathers were categorised by level of contact, and urban and rural differences were examined. Population mixing: high contact paternal occupations by urban/rural area	For those aged 0-4 years in rural counties in the highest contact group, there was an elevated risk of developing leukaemia (OR = 3.47). In rural counties there was a significant positive trend across the contact groups (no trend for urban areas).
Parslow et al (2002)	0-14 yrs Yorkshire England 1986-1996	Ecological study. Incidence rate ratios and population mixing levels were calculated for electoral wards and Poisson regression analysis conducted. Population mixing: diversity of origin of in-migrants (Shannon Index of Diversity) child and any age	Incidence of childhood leukaemias were significantly lower in areas of high child migrant diversity (IRR = 0.72) after controlling for population density and deprivation. There was also a non-significant but increased IRR in the lowest decile of child migrant diversity.
Law at al (2003)	0-14 yrs England Scotland & Wales 1991-1996	Case-control. Conditional logistic regression models were fitted to data for cases and their controls. Univariate and multivariate models including deprivation and population mixing variables were created. Population mixing: the volume of people who moved & the diversity of their origins (Shannon Index of Diversity)	In the ALL group, the odds ratio was significantly raised for the lowest category of diversity in all-age population mixing (OR = 1.37, CI = 1.00-1.86). There was a marginally significant increased risk of ALL with rural density.

Study	Population	Methods & Population Mixing Measure	Results
Labar et al (2004)	0-14 yrs Croatia 1986-1999	Ecological study. Incidence of childhood leukaemia in counties potentially exposed to depleted uranium, chemical war damage or population mixing during the war period were statistically compared. Population mixing: in-migration of war refugees	Incidence of ALL was highest in four counties where population mixing had occurred during the war period ($P<0.05$). No significant increases in leukaemia incidence were noted in counties that potentially experienced chemical war damage or exposure to depleted uranium.
Pearce et al (2004)	0-25 yrs Northern England 1968-1997	Case-control study. Odds ratios and confidence intervals were estimated using conditional logistic regression for different paternal occupational contact groups. Population mixing: high paternal occupational contacts at birth	Increased risk of leukaemia & NHL in children whose father's occupational contacts were high or very high, compared to the standard, and was most pronounced for ALL in ages 2-5yrs.
Wartenberg et al (2004)	0-14 yrs Rural areas USA 1973-1999	Ecological study. For 10 year periods, incidence rates were compared in rural counties that experienced varying degrees of population change, using logistic and Poisson regression. Population mixing: population growth in rural counties	For the period 1980 to 1989, the relative risk of ALL increased as a function of population change, from 1.9 in areas which noted up to 10% increase, to 2.6 in areas with a greater than 20% increase.
Feltbower et al (2005)	0-14 yrs Yorkshire England 1986-1998	Ecological study. Standardised incidence ratios were compared using Poisson regression to determine whether associations exist between leukaemia and type 1 diabetes incidence and population mixing. Bayesian methods of spatial correlation between diseases were also employed by modelling a bivariate outcome based on their standardized incidence ratios. Population mixing: The Shannon Index of Diversity	Higher rates of ALL (and type 1 diabetes) were present in areas of low migrant diversity even after adjusting for deprivation and population density. In areas with very high migrant diversity, significantly lower rates of ALL were observed (not for type 1 diabetes).
Rudant et al (2006)	0-7 yrs Communes France 1990-1998	Ecological study. Expected and observed numbers of leukaemia cases were compared by place of birth, population mixing, density and relative isolation. Population mixing: volume of in-migrants from another commune, another department and another region.	Increased risk of ALL with higher levels of in-migration for children residing in isolated communes at birth, with a population density of >50 people per km^2 (SIR = 2.59, CI = 1.48-4.49). Higher ALL with greater average distance travelled by in-migrants.
Kinlen (2006)	1-14 yrs West Cumbria England 1940-1943	Ecological study. Expected and observed deaths were compared for two areas in West Cumbria (& total for WC) where population mixing increased due to Royal Ordnance Factories (ROF) Population mixing: population growth due to ROF	Excess leukaemia deaths were found in 1-14 year olds during the construction and operation overlap phase (highest population mixing phase) with an observed/expected of 4.5 (CI = 1.1-12.2) and especially for 1-4 year olds (7.1, CI = 1.2-23.6).
Nyari et al (2006)	0-5 yrs South Hungary 1981–1997	Ecological study. Poisson regression was used to investigate the relationship between the risk of ALL and population mixing at the time of birth in county districts. Population mixing: the proportion of all incomers to the county district; and the proportion of incomers under 5 years	The risk of ALL increased significantly with increasing population mixing around the time of birth (trend across the range RR = 2.1, CI = 1.02–4.44). This effect was more marked for boys (RR = 3.1, CI = 1.13–8.51).
Clark et al (2007)	0-19 yrs Ohio USA 1996-2000	Ecological study. Counties were divided into quartiles of population change and standardised incidence ratios for each stratification were compared to rates for the whole state Population mixing: population growth	Significantly higher rates for ALL were noted in counties experiencing >10% population change 1990-2000, especially for children aged 1-4years in counties with 10-20% growth.

Study	Population	Methods & Population Mixing Measure	Results
Bellec et al (2008)	0-14 yrs France 1990-2003	Ecological study. Poisson regression was used to examine the association between childhood leukaemia and measures of migration by communes stratified by isolation and population density. Population mixing: proportion of incomers from outside the department, outside the region or in a distant commune. Also considered the weighted average migration distance travelled to each commune	A positive association was found with the proportion of migrants who came from a distant place, particularly among children aged 0-4 years in 'isolated' communes at the time of diagnosis (RR = 1.4, CI = 1.1-1.8 in the highest category of migration distance).

RR = relative risk, OR = odds ratio, IRR = incidence rate ratio, SIR = standardised incidence ratio, CI = 95% confidence intervals

Table 2.2: Studies on population mixing and childhood type 1 diabetes

Study	Population	Methods & Population Mixing Measure	Results
Fear et al (1999)	0-14 yrs Yorkshire England & Northern Ireland 1993-1994 & 1990-1992	Two case-control studies. Logistic regression was used to estimate odds ratios and confidence intervals for the relationship between type 1 diabetes incidence and parental occupational mixing. Population mixing: high parental occupational contacts	Type 1 diabetes was not significantly associated with high levels of parental occupational mixing in either setting. However, mothers with high occupational mixing conferred a non-significant reduced risk among children diagnosed under 5 years of age (OR = 0.58, CI = 0.24-1.38).
Parslow et al (2001)	0-14 yrs Yorkshire England 1986-1994	Ecological study. Regression models were used to calculate the effect of any age and childhood population mixing on the incidence of childhood diabetes, controlling for population density, ethnicity & the proportion of migrants. Population mixing: diversity of origin of in-migrants (Shannon Index of Diversity) child and any age	Areas with low levels of population mixing (bottom decile), were associated with significantly higher incidence rates of childhood diabetes in the 0-14 years age group (IRR = 1.46). When stratified by age, different effects were observed with raised IRR for ages 5-9 (IRR = 2.23) and 10-14 (IRR = 1.47) and decreased IRR for 0-4 year olds (IRR = 0.56).
Feltbower et al (2005)	0-14 yrs Yorkshire England 1986-1998	Ecological study. Standardised incidence ratios were compared using Poisson regression to determine whether associations exist between leukaemia and type 1 diabetes incidence and population mixing. Population mixing: The Shannon Index of Diversity	Higher rates of type 1 diabetes (and ALL) were present in areas of low population mixing even after adjusting for deprivation and population density.
Cox (2007)	0-14yrs Tayside Scotland 1998-2001	Ecological study. Poisson regression was used to test associations with demographic, socioeconomic & population mixing measures. Population mixing: migrant diversity, % in-migrants, % immigrants, % people in overcrowded houses, population density and a categorical population mixing variable combining migrant diversity & % in-migrants. All variables calculated for all ages, and children separately.	Areas with a higher % of child in-migrants had significantly lower rates of type 1 diabetes (Coeff = -0.16). Also, areas with low child in-migration & low child migrant diversity had the highest incidence of type 1 diabetes (Coeff = 1.67).

OR = odds ratio, IRR = incidence rate ratio, Coeff = coefficient, CI = 95% confidence intervals

2.2.4 Summary of population mixing studies

From Table 2.1 and the selection of articles reviewed above, a complex picture emerges. In terms of childhood leukaemia, 27 out of the 31 studies reviewed show increases of childhood leukaemia to be associated with high or increased levels of population mixing. In the limited number of type 1 diabetes studies conducted to date (Table 2.2), three studies showed an increase in type 1 diabetes to be associated with areas of low population mixing, and one study showed a non-significant protective effect of high maternal occupational contacts.

There are a number of possible explanations for the varied results described above. The most likely, is that the different measures of population mixing utilised by these studies could account for this variation. The majority of the studies which utilise a relatively simple measure of population mixing (population growth or proportion of in-migrants in an area) have shown consistently high rates of childhood leukaemia associated with high/increased levels of population mixing. Research which has employed the more complex measure of the Shannon Index of Diversity is slightly more mixed, but the majority of studies show the opposite relationship. In six out of nine studies completed, low population mixing has been related to higher incidence of disease, two studies show no association, and only one study has recorded a positive trend. In addition, variations in the timing of the population mixing measurements within study periods are likely to be important. The majority of early leukaemia studies measured increases in population directly prior to the disease study period. In contrast, more recent studies which tend to cover larger geographical areas and rely on national census data for measuring population mixing, have paid less attention to the timing of these measurements. Moreover, recent studies have tended to test different variations of Kinlen's original hypothesis (or the delayed infection hypothesis for diabetes) and thus are not strictly equivalent to the early leukaemia work. Other possible explanations include the use of different geographical units of analysis, different lengths of study periods, different age groups of children, different diagnostic criteria, unexplained confounding, or simply that the studies are not directly comparable.

2.3 Conclusion

Despite the mixed findings of the most recent work on population mixing and health, there is now substantial evidence to support a role for population mixing in the aetiology of childhood leukaemia. Its role in childhood type 1 diabetes pathogenesis is also receiving increasing attention. The studies reviewed in this chapter have shown that a range of population

movements, varying over different temporal and spatial scales and consisting of different groups of individuals (e.g. children, international migrants), are important in disease causation.

However, a number of criticisms of the research to date can be identified. A major criticism relates to how population mixing has been defined in previous studies. Most studies have defined, and therefore measured, population mixing as population change. This measure does not accurately enumerate the volume of new people entering an area, and does not take into account where they originate. Furthermore, no studies have as yet examined the effect of shorter term, long distance movements, such as tourist flows. Such movements have the potential to introduce local children to an increasing range of new infections, as tourists become more widespread within previously small communities. In addition, some of the most recent studies have tended to use static measures of population mixing (e.g. high levels of migration/diversity) at fixed points in time, with little concern for the timing of these measurements. The original hypothesis clearly stated that where levels of population had *increased* over time (usually over short time periods), excesses of childhood leukaemia could be expected in the years that follow. Thus, there is a need to assess how levels of population mixing have changed over time, and also for greater attention to be paid to the timing of the measurements, whilst routinely accounting for the latency period of the diseases. Finally, this topic has received little attention in New Zealand. As noted in Chapter 1, migration and population mixing are increasingly important facets of everyday life in this country. Only one rather limited study on childhood leukaemia has thus far been achieved in New Zealand (Dockerty et al., 1996a), leaving a research gap which requires attention. The following two chapters set out the basic epidemiology and aetiologies of childhood ALL and type 1 diabetes, with specific reference to infections and their links with population mixing. The importance of examining these diseases in the New Zealand setting is also considered.

Chapter 3: Childhood acute lymphoblastic leukaemia

3.1 Introduction

Although overall incidence is rare, leukaemia is the most common paediatric cancer in affluent societies (Greaves, 2002, Parkin et al., 2003). It has been estimated that around 49,000 new cases of childhood leukaemia were diagnosed worldwide in 2000 (World Health Organization, 2007). Incidence of the disease is thought to be increasing in many parts of Europe (Shah and Coleman, 2007, Steliarova-Foucher et al., 2005, Steliarova-Foucher et al., 2004), America (Xie et al., 2003), Australia (Milne et al., 2007) and New Zealand (Dockerty et al., 1996b). However, the reasons for the increase, and the precise aetiology of the disease, remain debated (Terracini and Maule, 2007). This chapter summarises the current epidemiological and aetiological knowledge regarding childhood leukaemia, with a specific focus on acute lymphoblastic leukaemia.

3.2 Acute lymphoblastic leukaemia

Acute lymphoblastic leukaemia (ALL) is a distinct form of leukaemia which affects the white blood cells in the bone marrow (Leukaemia Foundation, 2003), and is the predominant subtype of this malignancy found in children (Parkin et al., 2003). In the mainly white populations of North America, Oceania and Europe, approximately 80 percent of leukaemia patients have ALL (Little, 1999). ALL is characterised by the accumulation of immature white blood cells in the bone marrow which replace normal bone marrow tissue (Liesner and Goldstone, 1997). Without treatment the bone marrow will produce increasing numbers of these abnormal cells and the production of normal white blood cells (vital for fighting infection) will fail almost completely (Redaelli et al., 2005). The aberrant cells can eventually spill into the blood and are then carried throughout the body. Symptoms of this disease include fatigue, weight loss, repeated infections and excessive bruising and nose bleeds (Ruddon, 2007). If the disease is left untreated it becomes fatal fairly rapidly (Hoffbrand et al., 2001). With treatment, survival rates in children have increased steadily since the 1960s to around 80 percent (Pui et al., 2004). However, most current treatments are biologically crude and are often associated with considerable toxicity and morbidity (Greaves and Wiemels, 2003). In addition, a recent study suggests that incidence of secondary neoplasms increase steadily over 30 years after successful treatment of ALL (Hijiya et al., 2007).

3.3 Age, sex and ethnicity

In developed countries, peak incidence of childhood ALL usually occurs between one and four years of age (Little, 1999), followed by falling rates in later childhood and adolescence (Liang and Pui, 2005). This peak has not always been present and was first apparent in mortality data in England and Wales in the 1920s, and became more pronounced by the 1940s (Hewitt, 1955). Similar peaks were noted in white Americans in the 1940s (Gilliam and Walter, 1958). However, the incidence peaks did not emerge in black Americans (Ross et al., 1994) and Japan (Ajiki et al., 1994) until the 1960s. In developing countries in South America and Asia the childhood peak is less marked, and in sub-Saharan Africa there is little or no sign of a peak in early childhood (Stiller and Parkin, 1996). In New Zealand a marked peak in incidence was observed at ages 2-3 years throughout the period 1968 to 1990. The early age peak was prominent for both Māori and non-Māori children. However, incidence of ALL was significantly lower in Māori compared to non-Māori children (Dockerty et al., 1996b). In general, incidence has been found to be higher in children of European descent compared to those of African descent (Liang and Pui, 2005). For example, in Los Angeles in the USA, Hispanic children had the highest rates of ALL (48.0 per million) for the period 1984-1992, followed by non-Hispanic white children (39.2 per million) and then black children (16.1 per million) (Parkin et al., 1998).

Incidence of ALL is almost universally higher in boys than girls (Liang and Pui, 2005). Only three out of 47 countries around the world reported a slight female excess in ALL cases during the 1980s and early 1990s (Parkin et al., 1998). Data in New Zealand for 1968-1990 also revealed a statistically significant male excess in ALL cases in children aged 0-14 years (age-adjusted relative risk for male/female cases = 1.36) (Dockerty et al., 1996b). In England and Wales, incidence and mortality were at least 15 percent higher in boys than girls throughout the twentieth century (Shah and Coleman, 2007).

3.4 Temporal trends

A number of studies have examined temporal patterns in childhood leukaemia in recent decades, with the majority showing that rates of the disease have increased (Maule et al., 2006). For example, data from 63 European population-based cancer registries revealed that lymphoid leukaemia rose by an average of 1.4 percent per year between 1970 and 1999 (p-value <0.0001) (Steliarova-Foucher et al., 2004). A recent study in England and Wales showed that the incidence of childhood leukaemia rose by 20 percent between the early 1970s and 2000. Taken

together with increases in mortality found between 1911 and 1950, these data suggest a century long increase in the incidence of the disease (Shah and Coleman, 2007). In the USA, the overall incidence rate of leukaemia increased significantly (estimated annual increase of 0.5 percent) in those aged 0-20 years for the period 1973-1998 (Xie et al., 2003). Similar, but non-significant trends have been noted in Australia. A weak upward trend in incidence (of 0.40 percent per year) was observed in Western Australia between 1960 and 2006. Data for the whole of Australia were also consistent with a weak increase in the incidence of childhood ALL since 1982 (Milne et al., 2007). However, in other settings there is contrary evidence. No significant change was seen in the overall incidence rate for acute lymphoblastic leukaemia between 1982 and 2001 in the Nordic countries (Hjalgrim et al., 2003), nor in France for the period 1990 to 1999 (Desandes et al., 2004). The most recent data available for New Zealand revealed a significant increase in age-standardised incidence rates of ALL between 1968 and 1990 (p-value = 0.02). During 1968-1972 the rate per 100,000 person years was 3.22, compared to 4.10 per 100,000 person years by 1988-1990 (Dockerty et al., 1996b).

3.5 Geography

According to the most recent figures available from the International Agency for Research on Cancer (IARC), the incidence of childhood leukaemia varies substantially at a number of geographical scales. At the continent-level, higher incidence has been observed in North America and Europe compared to Asia and Africa. Furthermore, considerable variations in incidence have been noted within continents. For example, in Europe age-standardised incidence rates per million ranged from 18.5 in Bulgaria to 42.8 in Denmark (Parkin et al., 1998). A recent and more detailed study of childhood leukaemia incidence in Europe supports these findings. The authors found that the average annual age-standardised rate for childhood leukaemias (1970-1999) was 39.3 per million in eastern Europe, compared to 45.7 in western Europe. Higher incidence in the west was mainly due to a more pronounced peak in the occurrence of lymphoid leukaemia at age 2-3 years (Steliarova-Foucher et al., 2004).

Large differences in incidence have also been reported between countries. For the period 1980-1992 a ten-fold variation in childhood leukaemia incidence was noted between countries. Namibia in Africa had the lowest age-standardised incidence rate of 4.5 cases per million compared to the highest of 46.3 cases per million in Costa Rica (Figure 3.1). For approximately the same time period, age-standardised incidence in New Zealand non-Māori was 41.3 cases per million compared to 21.9 cases per million in Māori children (Parkin et al., 1998). Within

countries, a significant geographical variation in incidence was found between the municipalities of Sweden (Samuelsson and Lofman, 2004). Furthermore, a study in Yorkshire in the UK has revealed extensive spatial variation in the incidence of childhood ALL for even smaller areas (electoral wards, which in 1991 had a median childhood population count of 700) (Feltbower et al., 2005). There are currently no published data on how childhood leukaemia varies geographically within New Zealand.

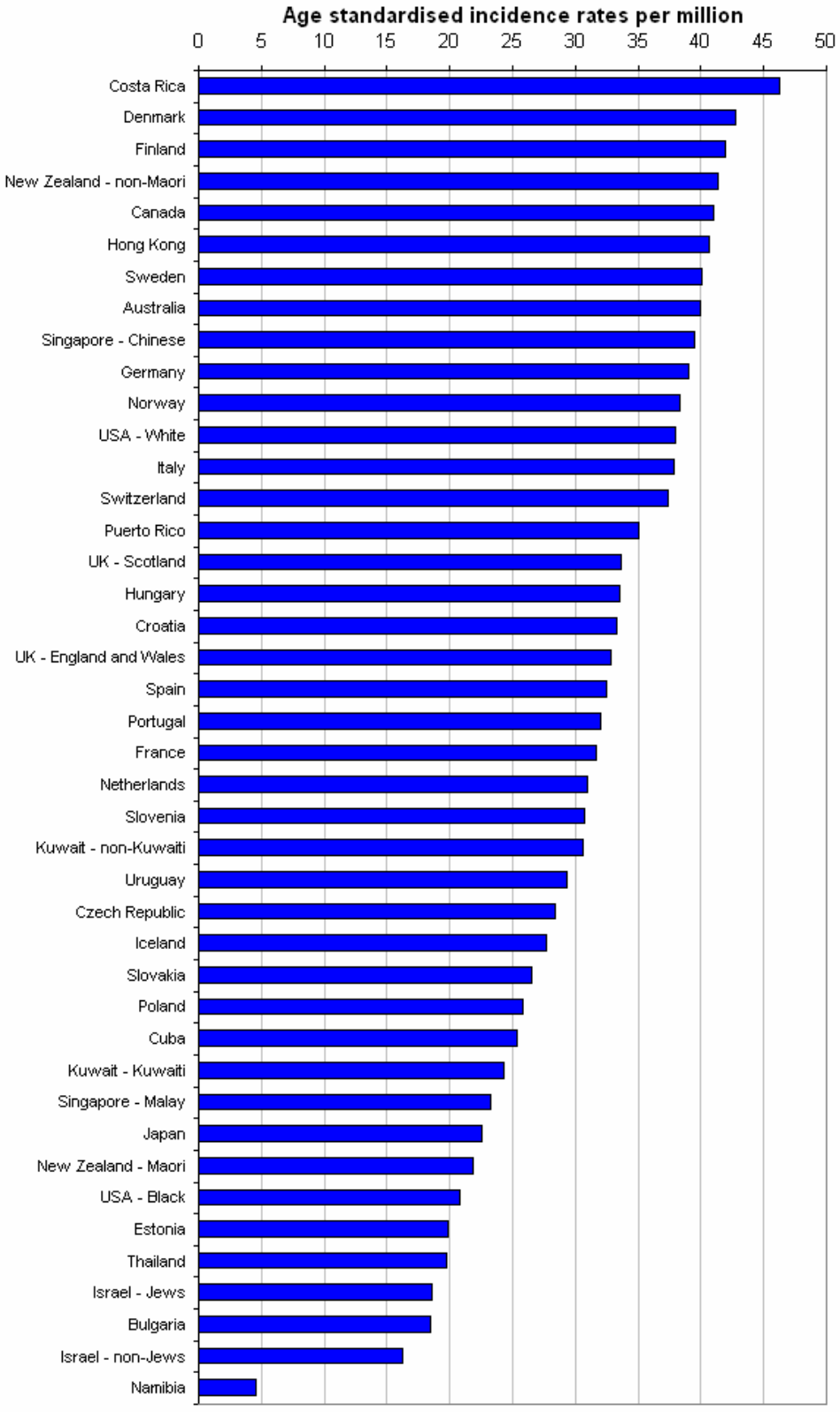

Figure 3.1: Lymphoid leukaemia incidence per million by country 1980-1992
(data from: Parkin et al 1998)

3.6 Aetiology of childhood leukaemia

Although the incidence of childhood leukaemia has been shown to vary both geographically and temporally, the reasons for these differences are unclear. Moreover, despite a plethora of research on the topic, the causes of childhood leukaemia remain enigmatic. However, it is thought that the disease arises from a combination of genetic susceptibility and environmental exposures, both of which may partially explain the geographical and temporal patterns in incidence of the disease (Belson et al., 2007). The following section will firstly summarise the current knowledge regarding the genetic aspects of childhood leukaemia, before reviewing a selection of possible environmental risk factors for the disease.

3.6.1 Genetic susceptibility

The precise role of genetic susceptibility in the aetiology of childhood leukaemia is not yet known (Heath, 1996). However, children with certain hereditary syndromes are much more susceptible to this disease. For example, children with Down's syndrome have a 10- to 20-fold increased risk of developing ALL (Liang and Pui, 2005). Other hereditary diseases associated with an increased incidence of childhood leukaemia include Bloom's syndrome, Fanconi anemia, ataxia, Shwachman syndrome and neurofibromatosis. These diseases are characterised by defective DNA repair, an abnormal number of chromosomes, or chromosomal abnormalities (Belson et al., 2007). However, such conditions are involved in relatively few cases (Stiller, 2004).

Familial inheritance has also been noted, especially in siblings and twins (Little, 1999). Siblings of affected patients have a 2- to 4-fold increased risk of developing ALL than children whose siblings do not have the disease (Draper et al., 1977, Redaelli et al., 2005). Risks of a similar magnitude have been noted for non-identical (dizygotic) twins (Liang and Pui, 2005). Precise calculation of the concordance rates in identical (monozygotic) twins is problematic due to the rarity of the condition, the limited availability of accurate databases, and possible ascertainment biases. A review of the published data suggests that in 10-15 percent of identical twin pairs where one twin has ALL, the other will also develop this disease (Greaves et al., 2003). The familial aggregation of ALL may result from an inherited predisposition or shared environmental factors (Bhatia and Robison, 1999). Molecular studies have demonstrated that concordance in identical twins is in fact not genetic, and instead reflects a shared intrauterine blood circulation by which leukaemia arising in one twin is spread to the co-twin (Greaves et al., 2003, Liang and

Pui, 2005). Although some degree of inherited susceptibility is very likely to exist, it has been estimated that a strong inherited predisposition is only important in a small minority (approximately 5 percent) of childhood leukaemia cases (Greaves, 2006).

Whatever the role for inheritance of ALL susceptibility, it is now known that the main subtypes of ALL involve a large variety of genetic alterations and are characterized by chromosome changes or translocations. Chromosome translocations are the illegitimate recombination or 'fusion' of normally separate genes. *TEL–AML1* is the most common fusion gene in paediatric cancers and is present in 22 percent of patients with ALL (Shurtleff et al., 1995). There is currently no epidemiological evidence identifying candidate *in utero* exposures that might induce the DNA breaks that are thought to be necessary for *TEL–AML1* fusion. These chromosome abnormalities might arise spontaneously (and frequently) as accidental by-products of haematopoiesis (blood formation) (Greaves and Wiemels, 2003). However, leukaemia is a relatively rare disease; the cumulative risk of any child developing leukaemia before the age of 15 years is one in 2,000 (Parkin et al., 1998). According to transgenic experiments in mice, a single chromosome translocation is not usually enough to generate leukaemia. Furthermore, a recent human study has revealed that common leukaemia fusion genes are present in the cord blood of healthy newborn babies at a frequency that is 100-fold greater than the risk of developing leukaemia (Mori et al., 2002). In other words, pre-leukaemic cells are present in the cord blood of many more babies than actually acquire leukaemia. Thus while the necessary chromosome translocations may be common, an additional postnatal 'event' is clearly required for the development of overt leukaemia.

3.6.2 Environmental factors

It is now widely believed that the final postnatal event required to prompt the onset of childhood leukaemia, is environmental (Greaves, 2006, McNally and Parker, 2006). While many candidate environmental triggers of this disease have been suggested (McNally and Parker, 2006), this review will focus on some of the more developed research: the role of radiation (both ionising and non-ionising), chemicals, and infections.

3.6.2.1 Radiation

Ionising radiation

The only established causal risk factor in childhood leukaemia is ionising radiation (Belson et al., 2007, Greaves, 2006). Research of individuals episodically exposed to radiation due to atomic bombs, occupational exposure and medical treatment form the basis of evidence for this association. Moreover, animal models offer support for the role of radiation in leukaemia pathogenesis in mice (Anderson, 2006b, Anderson et al., 2000).

Using data from the Life Span Study cohort of atomic bomb survivors in Japan during the late 1950s to 1987, it has been shown that radiation induces risks for all subtypes of the disease (Preston et al., 1994). In addition, a large case-control study conducted in the UK in the 1950s found that radiography of a mother's abdomen during pregnancy increased the risk of leukaemia in her baby by around 50 percent (Stewart et al., 1958). A recent meta-analysis corroborated this result and found evidence of a dose-response relationship between the number of x-ray films used in the examination and the relative risk of cancer. Furthermore, a reduction in relative risk over time corresponds with a reduction in foetal radiation dose, and the results of animal studies show foetuses to be susceptible to cancer induced by radiation (Doll and Wakeford, 1997). As a result of this discovery, virtually no pregnant women have abdominal radiography nowadays (Dickinson, 2005).

In addition, natural background ionising radiation has also been implicated in causing the disease. The largest component of exposure to natural background radiation in most countries is residential exposure to radon and its decay products (Wakeford, 2004). Based on a linear extrapolation of cancer risks from intermediate to very low doses (Brenner et al., 2003), it has been estimated that natural radiation could be responsible for one-fifth of childhood leukaemia cases in Britain (Wakeford, 2004). However, recent case-control studies do not support this model. No association was found between higher radon (UKCCS Investigators, 2002a) or gamma radiation (UKCCS Investigators, 2002b) levels in the home around the time of cancer diagnosis in children in the UK. This finding is consistent with case-control studies conducted in the USA (Lubin et al., 1998) and in Lower Saxony in Germany (Kaletsch et al., 1999). Radon levels in New Zealand are thought to be low compared to those reported in other countries (National Radiation Laboratory, 1998, Robertson et al., 1988).

Increased risk of leukaemia has also been associated with proximity to and employment in, the nuclear industry. In the early 1980s, excess cases of childhood leukaemia were reported around the nuclear reprocessing plants of Sellafield (Black and Independent Advisory Group, 1984, Committee on Medical Aspects of Radiation in the Environment, 1986) and Dounreay in the UK (Committee on Medical Aspects of Radiation in the Environment, 1988), and around La Hague in France (Guizard et al., 2001, Viel et al., 1995). However, radiation doses from environmental exposure to discharged radioactivity were generally found to be less than the doses received from natural background radiation, and were thus much too low to account for the excess leukaemia cases (Committee on Medical Aspects of Radiation in the Environment, 1996, Rommens et al., 2000). A controversial case-control study of children and their parents living in West Cumbria between 1950 and 1985 indicated that the Sellafield cluster of cancer cases could be explained by preconceptional exposure to radiation of fathers employed at the plant. Significantly higher relative risks were found for children of fathers employed at Sellafield at their conception, and this risk was especially high where fathers received a total preconceptional ionising radiation dose of 100 mSv or more (Gardner, 1992, Gardner et al., 1990). However to date, this association has not been corroborated in studies of the offspring of Japanese atomic bomb survivors, many of whom received much higher radiation doses than the Sellafield workers (Izumi et al., 2003, Yoshimoto et al., 1990). Furthermore, paternal occupation at other nuclear plants could not account for similar observed excesses (Committee on Medical Aspects of Radiation in the Environment, 2002, Pobel and Viel, 1997, Urquhart et al., 1991). As a result, the theory that paternal preconceptional irradiation can trigger childhood leukaemia is now considered unlikely (Doll et al., 1994, Wakeford, 2004).

New Zealand has no nuclear power stations or reprocessing plants. However, fallout deposition from nuclear weapons testing conducted in the Pacific occurred in New Zealand between 1953 and 1975 as a result of tests by the USA, USSR, UK, France, and to a lesser extent, China. Heaviest deposition occurred in 1964 when an average of 130 Bq of strontium-90 and 250 Bq of caesium-137 per square metre were deposited across all New Zealand monitoring sites. Concentrations of strontium-90 and caesium-137 in cows milk peaked in 1965 and decreased to almost undetectable levels by 1986 (Matthews, 1993). If nuclear fallout was important in explaining childhood leukaemia cases, a peak in incidence would be expected shortly after 1965, followed by a decline. However, a continued increase in childhood leukaemia incidence has been noted between 1953 and 1990, and thus is unlikely to be related to radioactive fallout (Dockerty et al., 1996b).

Non-ionising radiation

Forms of non-ionising radiation have also been implicated in the aetiology of childhood leukaemia. For over 20 years extremely low frequency electromagnetic fields (EMFs), produced by alternating electric currents such as those found in high voltage power lines, have been suspected of increasing the risk of the disease (Dickinson, 2005). However, the evidence to support this association is considered weak (Greaves, 2006). A pooled analysis from nine studies found no increased risk of childhood leukaemia in magnetic fields averaging under 0.4 μT. Nevertheless, for the small (0.8) percentage of children who lived in homes with exposure levels greater than 0.4 μT, the risk of leukaemia doubled (Ahlbom et al., 2000). This finding was supported by a population-based case-control study in Germany which found a significant positive relationship between childhood leukaemia risk and magnetic field exposure at night (Schüz et al., 2001a). However, as the authors note, selection bias could partly explain these results. Furthermore, even if the association is real, exposure to levels of greater than 0.4 μT is rare. Other studies have found little or no effect for direct exposure to magnetic fields (Schüz et al., 2001b, UKCCS Investigators, 2000) or electric fields (McBride et al., 1999, UKCCS Investigators et al., 2002). Furthermore, laboratory experiments have provided no good evidence that extremely low frequency electromagnetic fields are capable of producing cancer (Advisory Group on non-ionising radiation, 2001). For a comprehensive review see McNally and Parker (2006) and Wakeford (2004). A recent nationwide study of childhood leukaemia and EMF exposures in New Zealand found raised but non-significant odds ratios for the highest category of magnetic field exposure in the child's bedroom (Dockerty et al., 1998, Dockerty et al., 1999a), and a time-weighted average of their bedroom and day room (Dockerty et al., 1999b). However, as the authors note, the small number of cases (4) in the highest exposure category and the fact that multiple comparisons were made, mean that the results of this study are equivocal. According to the EMF measurements recorded by this research, few houses had mean magnetic fields above 0.2 μT (Dockerty et al., 1998).

3.6.2.2 Chemicals

As with radiation sources, the potential aetiological role of chemical exposures have been investigated and received much public interest. The chemicals most commonly associated with childhood leukaemia are hydrocarbons and pesticides. Studies of childhood exposure to these chemicals have considered both direct exposure (e.g. use around the home) and indirect exposures (e.g. parents occupational exposures) (Belson et al., 2007). The most widely

recognised hydrocarbon is benzene, which is a naturally occurring organic compound. It can be found in crude oil (and therefore petrol) and has been widely used in the manufacture of paints, paint removers, plastics and adhesives. Benzene has long been recognised as a genotoxic carcinogen with the strongest evidence linking it to lymphohaematopoietic cancers (Duarte-Davidson et al., 2001). It has a strong dose-response relationship with leukaemia, particularly acute-myeloid leukaemia (Belson et al., 2007). Numerous studies have described the adverse health effects of benzene, with the majority focusing on adult occupational exposure (Duarte-Davidson et al., 2001). A recent study found an increased risk of leukaemia to be associated with cumulative benzene exposures at lower levels than had previously been reported (Glass et al., 2003). However, studies of childhood benzene exposure show mixed results. For instance, a case-control study in Italy estimated exposure to benzene from traffic exhaust using data on traffic density (vehicles/day) on nearby main roads, distance between roads and residence, and information on vehicle emissions and weather conditions. This study found a significantly increased risk of leukaemia in children who were heavily exposed to benzene (relative risk = 3.91, 95 percent confidence intervals = 1.36-11.27) compared to children whose homes were not exposed to road traffic emissions (Crosignani et al., 2004). A recent study in France also noted a positive association between childhood leukaemia and living next to a petrol station or repair garage, which are known benzene emitters (Steffen et al., 2004). In contrast, a large case–control study conducted in Denmark, using a validated model to estimate lifetime exposure to benzene from motor vehicle exhaust, did not find an association with childhood leukaemia (Raaschou-Nielsen et al., 2001). Studies in the US have also failed to find an association between traffic density and childhood leukaemia (Langholz et al., 2002, Reynolds et al., 2001, Reynolds et al., 2002). However, a more recent study found an elevated risk and a significant trend for childhood leukaemia with increasing exposure to hazardous air pollutants in general (Reynolds et al., 2003).

Many studies have suggested a link between pesticide exposure and childhood leukaemia (Belson et al., 2007). Children are most frequently exposed to pesticides in the home and garden (Grossman, 1995). However, other sources include local agricultural applications, parental occupation, pet products and contaminated food (Belson et al., 2007). A critical review of studies published between 1970 and 1996 found that in general, the relative risk estimates were modest and were stronger when pesticide exposure was measured in detail. Frequent occupational exposure to pesticides or home pesticide use was more strongly associated with childhood leukaemia than professional exterminations and garden pesticide use (Daniels et al., 1997). However, a study of children with ALL in Canada found that indoor use of some insecticides,

and pesticide use in the garden and on interior plants, were associated with increased risks of up to several-fold in magnitude, especially with frequent use in the prenatal period (Infante-Rivard et al., 1999). Household pesticide use was also examined in the Northern California Childhood Leukaemia Study. The results showed that the use of professional pest control services at any time from one year before birth to three years after birth significantly increased the risk of childhood leukaemia. Furthermore, significantly raised odds ratios were noted for exposure to insecticides three months before pregnancy, during pregnancy and in the first and second year after birth (Ma et al., 2002a). The results of a case-control study in Cumbria in England did not support a role for preconception paternal occupational exposures to pesticides or herbicides in the aetiology of childhood lymphoid leukaemia (Pearce et al., 2006c).

A recent review evaluated the research on this topic conducted between 1998 and 2004 (Nasterlack, 2006). Collectively, the studies were found to be supportive of an increase in the risk of different cancer types associated with exposure to pesticides. For childhood leukaemia, five studies found a positive association with some kind of pesticide exposure, whereas no association was noted for eleven other exposures. An update revealed that most, although not all, of the case-control studies published between 1998 and 2006 reported an elevated leukaemia risk associated with at least one of a large variety of exposure categories. However, the highest risks noted had the most imprecise risk estimates. Reviewing cohort and ecological studies found that all of the reported risks were around unity or below unity (Nasterlack, 2007).

It should be noted that many of the studies examining the role of chemical exposures (to hydrocarbons or pesticides) in the aetiology of childhood leukaemia suffer from a lack of individual exposure data, potential confounders, a small number of exposed children, or possible recall bias. Further research addressing these design flaws is therefore necessary to elucidate any real associations with childhood leukaemia.

3.6.2.3 Infection

One potential risk factor, which has received substantial attention in the literature and which has been investigated using many different study designs and methods, is infection. An infective basis for childhood leukaemia has been suspected since the early twentieth century (Cooke, 1942, Ward, 1917). The rationale for this view included a similar peak age distribution for childhood leukaemia and common infectious diseases, such as diphtheria and measles (Pierce, 1936). Furthermore, many leukaemia patients were found to have a record of infection prior to,

or coincident with, diagnosis (Greaves, 2006). Infection as a cause lost favour when it was recognised that leukaemia was not contagious (Ward, 1917) but infection is now being reconsidered due to the discovery of viruses as causative factors in adult T-cell leukaemia (Gallo et al., 1984) and in other human cancers (e.g. Epstein-Barr virus in Burkitt lymphoma) (Belson et al., 2007). The finding that viruses cause leukaemia in domestic cattle, cats and chickens has also helped substantiate this theory (Greaves, 2006, Jarrett, 1987, Penrose, 1970). In the late 1980s two main hypotheses were proposed to explain the role of infections in causing childhood leukaemia: the population mixing hypothesis; and the delayed infection hypothesis.

Infection hypotheses

Leo Kinlen's population mixing hypothesis arose from the controversy surrounding the childhood leukaemia clusters observed around the Sellafield and Dounreay nuclear reprocessing plants in the UK (Kinlen, 1988, Kinlen and Doll, 2004). He, and others, doubted the credibility of claims that radiation was the cause of these clusters (Greaves, 2006). Kinlen's population mixing hypothesis was thus formulated as an alternative theory to explain the clusters, and consisted of three main elements. First, that an influx of population into isolated rural areas is conducive to epidemics of certain infections. Second, that Sellafield and Dounreay were extreme examples of geographic isolation and population influx. Third, that some unidentified virus (or viruses) can cause childhood leukaemia (Kinlen, 1988). He tested this theory by examining leukaemia mortality in Glenrothes, a new town in rural Scotland with no nuclear plants and which experienced a doubling of its population in the 1950s and 1960s. A significant excess of leukaemia deaths (7 observed, 1.5 expected) was found in those aged under 5 years for the period of greatest population growth. This work together with subsequent studies in the UK and elsewhere (see section 2.2.3 and below) provided a "firm basis" for suggesting that some childhood leukaemia clusters may be an unusual outcome of an unknown common infection, introduced to susceptible children through population mixing (Greaves, 2006 p.196).

At the same time, Mel Greaves proposed his 'delayed infection' hypothesis (Greaves, 1988). This model sought to explain the peak incidence of ALL at 2-5 years and also the geographical and temporal trends in the incidence of the disease. A central tenet of this hypothesis is that the immune system is programmed to anticipate infectious exposure neonatally and in infancy. Absence of infections in early life (a feature of modern affluent societies), thus leaves the immune system weak, and subsequent exposure to infection can precipitate ALL. Therefore, according to the minimal two-hit model he proposed for leukaemia development, the first

mutation would occur *in utero* forming a pre-leukaemic clone, and the second mutation would be triggered some years later, probably by delayed infection (Greaves, 2006). Recent molecular evidence has corroborated this theory (Dickinson, 2005). Pre-leukaemic clones have been found in blood samples of new born babies who later developed ALL (Gale et al., 1997, Wiemels et al., 1999), suggesting an *in utero* initial event. Furthermore, many more babies have these pre-leukaemic cells in their cord blood than actually develop leukaemia, highlighting the need for post-natal events (Mori et al., 2002). Greaves' hypothesis follows a similar immunological argument to the 'hygiene hypothesis' which was initially formulated to explain childhood allergies (Strachan, 1989) and has since been applied to autoimmune diseases such as multiple sclerosis (Alvord et al., 1987, Bach, 2005a) and type 1 diabetes (Filippi and von Herrath, 2005, Gibbon et al., 1995, Kolb and Elliot, 1994) (see Chapter 4).

It is important to note that the hypotheses of Greaves and Kinlen are not mutually exclusive. They both posit that a period of low infectious exposure (through social or geographical isolation) in early life, followed by an abnormal response to later exposure to one or more common infections, is necessary to cause childhood leukaemia. Over the past 10 to 15 years evidence has accumulated to support these hypotheses (Dickinson, 2005). The remainder of this section will review this evidence.

Population mixing studies

As noted in the previous chapter, numerous studies have examined the association between population mixing and childhood leukaemia. Out of the 31 studies reviewed, 27 noted a significant excess in ALL or childhood leukaemia cases in areas/individuals exposed to high or increased levels of population mixing. At the area-level, Kinlen and colleagues have found significantly raised incidence of childhood leukaemia in rural new towns compared to overspill new towns (Kinlen, 1988, Kinlen et al., 1990), in areas which experienced increased commuting (Kinlen et al., 1991), in rural areas of Scotland which received a large influx of oil workers (Kinlen et al., 1993), and in Pembrokeshire near large rural construction sites, and in Scottish hydroelectric counties while construction was underway (Kinlen et al., 1995). Furthermore, raised childhood leukaemia has been noted after a number of unusual war-time population mixing events. For example, increased incidence of childhood leukaemia has been associated with rural areas of England and Wales which witnessed an influx of military personnel (Kinlen and Hudson, 1991) and with a high proportion of war-time children evacuees from London (Kinlen and John, 1994). In addition, an excess of leukaemia cases was found in Orkney and

Shetland when troops outnumbered local people during the second world war (Kinlen and Balkwill, 2001) and in West Cumbria when the construction and operation of Royal ordnance factories resulted in influxes of new workers to the area (Kinlen, 2006). Raised incidence of childhood leukaemia was also shown in other European countries with high levels of rural migration (Kinlen and Petridou, 1995). Individual-level analyses have revealed leukaemia incidence to be higher in children with fathers in very high contact occupations, especially those working in transport, in northern England (Pearce et al., 2004), in rural areas of Scotland (Kinlen and Bramald, 2001), in the UK as a whole (Kinlen, 1997), and in rural counties of Sweden (Kinlen et al., 2002).

While Kinlen's work has been criticised by various authors for focusing on unusual sociodemographic events, often occurring in small rural populations, and for using mortality data and crude changes in population size as a measure of population mixing (e.g. Parslow et al., 2005), the excesses found are so large and consistent across the different studies as to point to a genuine phenomenon (Doll, 1999, Greaves and Alexander, 1993). Moreover, work by other researchers conducted in increasingly diverse locations, overwhelmingly supports the findings of Kinlen and colleagues; that increased levels of population mixing are significantly and positively associated with childhood leukaemia. For example, ecological studies in Canada (Koushik et al., 2001), France (Bellec et al., 2008, Boutou et al., 2002, Rudant et al., 2006), Croatia (Labar et al., 2004), Hungary (Nyari et al., 2006), the USA (Clark et al., 2007, Wartenberg et al., 2004) and Hong Kong (Alexander et al., 1997), all found increased risk of childhood leukaemia in areas of high population mixing. Other studies in the UK also support these results (Dickinson et al., 2002, Dickinson and Parker, 1999, Langford, 1991, Stiller and Boyle, 1996).

However, four studies do not support these findings. Ecological work concentrated in Yorkshire in England between 1986 and 1996 found childhood leukaemia incidence to be significantly lower in areas of high population mixing, and significantly higher in areas of low population mixing (Parslow et al., 2002). These results were supported by further ecological work carried out in Yorkshire by Feltbower and colleagues (2005) and by a case-control study of children with leukaemia in England, Scotland and Wales (Law et al., 2003). The only study conducted in New Zealand found no significant relationship between age-adjusted rate ratios for childhood leukaemia in three population growth areas, when compared to the rest of New Zealand (Dockerty et al., 1996a). The inconsistencies in these findings may be explained by differences in how population mixing was measured, variable study design, differential control of confounding variables and in varying geographical and temporal scales of analysis.

Cluster studies

As well as being linked to the Sellafield and Dounreay leukaemia clusters in the UK, population mixing has also been associated with a number of pronounced clusters of childhood leukaemia in the US (Kinlen and Doll, 2004). The earliest example was observed between 1957 and 1960 in Niles, a suburb of Chicago in Illinois (Heath and Hasterlik, 1963). During this time, eight cases of ALL were observed in children aged 3-13 years, living in a single parish of Niles. The expected incidence for children aged less than 15 years was 1.6 cases. Seven of the eight affected families had children attending the same elementary school at the time of disease onset. Furthermore, the presentation of the leukaemia cases paralleled an excess of rheumatic illness in children from families using the elementary school. Researchers concluded that the cluster was probably a consequence of infection allied with the coincident six-fold population expansion that had occurred within the parish during the 1950s. Nearly all of the growth occurred in the northern part of the town which had previously been rural countryside (Heath, 2005, Heath and Hasterlik, 1963).

An even more striking cluster occurred recently in the small desert town of Fallon, in Churchill County, Nevada. From 1999 to 2001, 11 cases of childhood leukaemia were diagnosed, compared to less than one expected case. Ten of the observed cases were diagnosed with ALL. The age-standardized rate ratio in Churchill County as a whole was 12.0 (CI = 6.0-21.4, p-value = 4.3×10^{-9}). Researchers estimated that a cluster of this magnitude would be expected to occur in the US by chance once every 22,000 years (Steinmaus et al., 2004). An expert panel reviewing the potential causes of this cluster found no evidence to suggest that environmental contaminants (for example, arsenic levels in water and jet fuel emissions from the nearby airbase) were involved (Rubin et al., 2007, Sinks et al., 2004). The only possible cause which the panel could not exclude was the effect of a large increase of military personnel temporarily assigned to the Fallon Naval Air Station. In the early 1990s these personnel numbered 20,000, but reached 55,000 by 2000. However, the panel remained undecided as to whether this constituted an extreme population mixing event. Kinlen and Doll (2004) have since argued that the indirect exposure of Fallon in only a few years to such large numbers of people from outside the area represents "…a more extreme example of rural–urban population mixing than any of those traced and studied in Britain"(p.2).

The childhood leukaemia clusters at Sellafield, Dounreay, Niles and Fallon were investigated on a *post-hoc* basis making them particularly difficult to interpret. Numerous methods also exist for

examining clustering in areas without *a priori* assumptions as to whether or not they exist. For example, a substantial number of studies have tested whether childhood leukaemia has a tendency to cluster at both the time/place of birth and diagnosis. See Little (1999) for a comprehensive review of studies conducted prior to 1999, and see McNally and Eden (2004) for a more up to date review. Most recently, two studies carried out in the UK (Birch et al., 2000, McNally et al., 2006b) and one in Europe and Australia (Alexander et al., 1998), have revealed evidence of significant clustering of childhood leukaemia based on the time and place of diagnosis. Analyses carried out in Sweden and the UK (Gustafsson and Carstensen, 2000, McNally et al., 2002), have also found evidence for space-time clustering at the time and place of birth.

In New Zealand, two early studies (1953-1964) using the same data but different methods (Mantel and Knox) observed statistically significant space-time clustering of leukaemia in children aged 0-5 years at diagnosis (Glass et al., 1971, Gunz and Spears, 1968). For example, using the Mantel method, clusters were significant for additive constants of up to six months and six miles apart. However, the authors were unable to exclude the possibility that the clustering arose as a result of rapid population growth in New Zealand over the study period (Glass et al., 1971). A more recent study (1976-1987) was able to overcome some of the methodological limitations of the early studies, and found no significant clusters of childhood ALL overall. However, when the cases were analysed by age group at diagnosis, significant clustering was observed in children aged 10-14 years (p-value = 0.003). The authors concluded that this cluster could be a real or chance occurrence as several comparisons were made. If not due to chance, bias or confounding, the cluster could suggest a localised environmental risk factor or the person to person spread of an infectious agent (Dockerty et al., 1999c).

According to McNally et al (2006a), observing space-time clustering of ALL in general is indicative of an environmental component to its aetiology, and supports the involvement of infections in the pathogenesis of the disease. However, clustering would only be found if the infection were only to occur in short-lived epidemics or if it only affected a limited number of susceptible people (McNally and Eden, 2004). Furthermore, while being supportive of an infectious aetiology, the results of cluster analyses cannot directly test either Kinlen's or Greaves' hypotheses. Moreover, absence of spatial-temporal clustering should not be used as evidence for the absence of a particular risk factor (Law, 2005).

Individual-level studies

Other study designs have sought to examine the role of infections in childhood leukaemia pathogenesis. Due to the difficulties involved in conducting prospective research on such a rare disease, the majority of these studies are retrospective case-control investigations (Belson et al., 2007). Such studies have investigated maternal infection during pregnancy, early childhood infections (measured both directly or using proxy variables), and also the effects of infant breast-feeding.

In several case-control studies, infections acquired by mothers during pregnancy have been found to increase the risk of leukaemia in their offspring. For example, Epstein-Barr virus (Lehtinen et al., 2003) and lower genital tract infections in pregnant mothers (Naumburg et al., 2002), have been significantly and positively associated with childhood leukaemia. A raised, but non-significant odds ratio was reported for any infections during pregnancy in a case-control study in Scotland (McKinney et al., 1999a). However, recurring infections during pregnancy showed little effect on leukaemia outcomes in Canada (Infante-Rivard et al., 2000). Furthermore, in New Zealand, no relationship was found between childhood leukaemia and maternal infections during pregnancy or three months before conception (Dockerty et al., 1999d).

The results of studies examining childhood infections and subsequent leukaemia development are more mixed. A number of recent studies have noted a statistically significant increased risk in childhood leukaemia associated with exposure to infections (Chan et al., 2002, Dockerty et al., 1999d, Petridou et al., 2001, Roman et al., 2007). For example, a New Zealand study found an elevated risk of leukaemia in relation to reported influenza infections during the child's first year of life (Dockerty et al., 1999d). In the UK Childhood Cancer Study (UKCCS), children diagnosed with ALL had significantly more clinically diagnosed infectious episodes in infancy than controls, with the greatest difference noted in the neonatal period (less than one month) (Roman et al., 2007). However, in Greece the only positive association between infection and leukaemia was noted in children aged 5 years or older at diagnosis (Petridou et al., 2001). Several recent studies have also reported a statistically significant protective effect of early childhood infections (Chan et al., 2002, Jourdan-Da Silva and *al*., 2004, Ma et al., 2005, McKinney et al., 1999a, Neglia et al., 2000, Perrillat et al., 2002). Differences in study designs may help to explain the inconsistent results.

A further difficulty in attributing causality to specific infections is that the infection(s) relevant in the aetiology of leukaemia may not prompt obvious symptoms and thus could remain undiagnosed. In this situation the use of well known proxies for infectious exposure may be preferable (Greaves, 2006). Epidemiological studies have used a range of surrogate measures for estimating exposure to infections in early life. Most are related to the number of social contacts children encounter at young ages. For example, the number of siblings and birth order of case and control children have frequently been compared. Infections passed between siblings are thought to represent an important aspect of early immune system stimulation (Greaves, 2006). As a result, children with fewer siblings and a lower birth order would lack early infectious exposure from within the family, and would be expected to be at increased risk of developing childhood leukaemia. However, the results of recent case-control and cohort studies are varied (McNally and Eden, 2004). Of the 15 most recent studies, nine found no significant relationship with birth order (e.g. Gilham et al., 2005, McKinney et al., 1999a, Murray et al., 2002, Neglia et al., 2000, Perrillat et al., 2002), five found a significant increased risk associated with a high birth order (e.g. Bener et al., 2001, Jourdan-Da Silva and *al*., 2004, Shu et al., 2002), and one study found a significant protective effect of high birth order (Dockerty et al., 2001). A case-control study of ALL in Québec for the period 1989-1995 considered the importance of the number of siblings in the first year of life and at diagnosis. The results showed that having older siblings at the time of diagnosis significantly increased the risk of ALL in children diagnosed before four years of age, whereas having older siblings in the first year of life was protective among children diagnosed at age four or older (Infante-Rivard et al., 2000). These findings emphasise the importance of timing of exposure to infections and may help to explain the diverse results of studies conducted to date.

Another proxy often used as a measure of early life infectious exposure, is day-care attendance. It is well documented that use of day-care facilities increases the risk of childhood infections (Nystad et al., 1999). No studies have found an increased risk of childhood leukaemia associated with day-care attendance (McNally and Eden, 2004). However, significant protective effects have been noted in a number of case-control studies (Gilham et al., 2005, Infante-Rivard et al., 2000, Jourdan-Da Silva and *al*., 2004, Kamper-Jorgensen et al., 2008, Ma et al., 2002b, Perrillat et al., 2002, Petridou et al., 1993). Recently, the UKCCS revealed consistent reductions in the risk of ALL with increasing levels of social activity, for the period 1991 to 1996 (Gilham et al., 2005). A similarly large case-control study in Northern California (1995-1999) quantified exposure to other children and also reported a significant protective effect. Control children started day-care at a younger age, attended for a longer duration, remained in day-care for more

hours, and were exposed to more children at each day-care centre compared to children who later developed ALL (Ma et al., 2002b). A recent expansion of this study which included approximately twice as many ALL cases supported these results but only for non-Hispanic white children. There was no association between day-care attendance and ALL in Hispanic children (Ma et al., 2005). Other studies have found no significant association between day-care attendance and childhood leukaemia (Chan et al., 2002, Neglia et al., 2000, Rosenbaum et al., 2000). In New Zealand, Dockerty et al (1999d) found no significant relationship between regular contact with other children from outside the home during the first year of life and subsequent ALL development. As with all of the case-control studies reviewed in this section, a social class bias may exist between the case and control groups (McNally and Eden, 2004). Five studies have attempted to account for differences by socioeconomic status (Infante-Rivard et al., 2000, Ma et al., 2002b, Ma et al., 2005, Neglia et al., 2000, Perrillat et al., 2002), and all found a protective effect of early attendance at day-care facilities.

Breast-feeding is also thought to provide early exposure to infectious agents through close mother-to-child contact and also through passive antibody transfer (Kwan et al., 2004, McNally and Eden, 2004). Fourteen case-control studies examining the relationship between breast-feeding and childhood leukaemia were reviewed by the UKCCS investigators (2001). Four of the studies showed statistically significant reduced risks (Bener et al., 2001, Infante-Rivard et al., 2000, Shu et al., 1999, Smulevich et al., 1999), with the remainder revealing no significant associations (UKCCS Investigators, 2001). For example, no significant relationship was shown for children who were breast-fed compared to those who were not in a New Zealand study (Dockerty et al., 1999d). Combining the UKCCS findings with these other studies revealed an overall protective effect of being ever breast-fed compared to never breast-fed (UKCCS Investigators, 2001). A more recent study in Hungary revealed that breast-feeding for less than six months was associated with an increased risk of ALL when compared with feeding for longer than six months (Altinkaynak et al., 2006). Furthermore, a recent meta-analysis showed that both short-term and long-term breast-feeding reduced the risk of childhood ALL. Results for studies that adjusted and did not adjust for socioeconomic status were not significantly different from the results for the 14 studies combined (Kwan et al., 2004).

Mechanisms

Despite a large body of evidence implicating infections in the aetiology of ALL and childhood leukaemias in general, no individual species of viruses or bacteria have been identified (Greaves, 2006). However, several possible routes by which infection could act as a trigger for ALL have been proposed. First, a yet to be identified leukaemogenic virus could be involved which directly infects and transforms the beta precursor cells (Greaves and Alexander, 1993). Leukaemogenic viruses have already been identified in adult T-cell leukaemia (Gallo et al., 1984) and in leukaemia in domestic cats, cattle and chickens (Greaves, 2006, Jarrett, 1987, Penrose, 1970). However, it has been noted that the epidemiology of feline leukaemia is quite different to ALL in children, with a greater risk found early in life in crowded households (Anon, 1990). Second, proliferative stress could occur as a result of molecular mimicry. This process would involve a virus having certain features in common with a natural ligand (growth factor) in the body. Cell exposure to such a virus could mimic the proliferative effect of the ligand, and could provide a sustained stimulus. Third, infection of the bone marrow stromal cells (cells which produce bone, cartilage, fat, and fibrous connective tissue) by a virus could result in a dysregulation of growth factor production. Alternatively, systemic infection could induce an immune response which indirectly activates stromal cells to produce growth factors (Greaves and Alexander, 1993).

Infections summary

In summary, a large body of evidence now exists implicating infections in the aetiology of childhood leukaemia and ALL in particular. The overall consistency of the population mixing results together with evidence of spatial-temporal clustering of the disease, are suggestive of an important role for infections. Researching and measuring specific acquired infections is problematic due to the transient nature of many infections, and may be responsible for the varied results on maternal infections in pregnancy and child infections in early life. Of the proxy measures relating to infectious exposure, day-care attendance has been found to offer the most consistent protective effect. Breast-feeding studies have also noted a majority of inverse associations with childhood leukaemia, with the importance of the timing and duration of breast-feeding noted. The mixed nature of the results observed for sibling numbers also highlights the importance of timing in early social contacts and consequent childhood infections.

3.7 Conclusions

The most recent time trend data available for New Zealand revealed that childhood ALL increased significantly between the late 1960s and early 1990s. Similar increases have been noted in other affluent countries. The reasons for these increases and indeed the precise aetiology of the disease in general, remain unclear. However, it is currently thought that the pathogenesis of the disease involves damage to DNA before birth; possibly in response to infection, chemicals, ionising radiation, or other environmental exposures (McNally and Eden, 2004). These pre-leukaemic cells are then thought to be converted into overt disease after birth if children are susceptible (as a result of their genetic make up and early protection from infection) and if they experience one or more further events (Dickinson, 2005). Crucially in terms of this thesis, a delayed challenge from infections, possibly occurring after increases in population mixing, has been suggested as a potential final event.

Numerous studies in a variety of settings offer support for this theory. Kinlen and others have investigated all known examples of extreme urban to rural population mixing in the UK, and found significant positive results in each study. Work in Europe, North America, and Asia, corroborate these findings. In New Zealand, only one study has investigated the potential association between an increase in population mixing and childhood leukaemia incidence. However, as noted in Chapter 2, this study suffered from a number of limitations. In addition, studies in New Zealand are yet to examine the role of population mixing in the pathogenesis of childhood type 1 diabetes. As the following chapter notes, there is also evidence to implicate infections in the aetiology of childhood type 1 diabetes, and studies in the UK have found significant associations between this disease and area-level population mixing.

Chapter 4: Childhood type 1 diabetes

4.1 Introduction

Diabetes mellitus is one of the most frequent chronic diseases affecting children and adolescents (Chiarelli et al., 2005). In 2005 it was estimated that 430,000 children (0-14 years) had type 1 diabetes worldwide (International Diabetes Federation, 2005), with 1,056 children living with the disease in New Zealand (Wu et al., 2005). The global incidence of this disease is rising by around 3.4 percent per year but the reasons for this increase remain unclear (DIAMOND, 2006). This chapter examines the geographical epidemiology and aetiology of childhood type 1 diabetes. First, the chapter introduces the disease, describing what it is and who it affects. Second, the potential aetiology of the disease is reviewed with particular reference to environmental risk factors. Specific attention is paid to this disease in the New Zealand setting where data are available.

4.2 Type 1 diabetes

Type 1 diabetes is a chronic condition caused by the destruction of insulin producing beta cells in the pancreas. After the beta cells are destroyed, little or no insulin is produced by the body. Insulin is a hormone required to aid cell uptake of glucose, and without it, glucose accumulates in the blood resulting in diabetes mellitus (Feudtner, 2003). Two main forms of type 1 diabetes have been identified, type 1a which results from autoimmune processes, and type 1b which is thought not to be immune mediated (Rewers et al., 2005). Type 1a is the most common form, accounting for around 90 percent of type 1 diabetes cases in populations of European origin (Slama, 2003). Since the majority of type 1 diabetes patients in New Zealand are of European origin (Campbell-Stokes and Taylor, 2005, Ministry of Health, 1998), this thesis does not differentiate between the two types and type 1 diabetes is generally considered as an autoimmune phenomenon.

Every year, around 65,000 children aged under 15 years develop type 1 diabetes worldwide (International Diabetes Federation, 2005). These children present with a variety of symptoms, including weight loss, fatigue, polyuria (abundant urine), polydipsia (intense thirst), polyphagia (increased appetite), hyperglycaemia (elevated blood glucose) and ketoacidosis (elevated blood acids). In cases where insulin therapy is not initiated at this stage, the symptomatology worsens and the toxic acids can quickly poison the blood and induce a coma, and even cause death

(Feudtner, 2003, Slama, 2003). Since the discovery of insulin in 1922, death at onset is rare, but type 1 diabetes is still associated with many long-term health complications, for example retinopathy, nephropathy, neuropathy, accelerated atherosclerosis, coronary heart disease and stroke (Heller et al., In press, Slama, 2003). As a result, type 1 diabetes patients have a two- to ten-fold increased risk of mortality before 40 years of age (Karvonen et al., 2003). In Canterbury, New Zealand, a diagnosis of type 1 diabetes prior to 30 years of age is associated with a three-fold increase in mortality compared with the non-affected background population (Darlow et al., 2004).

4.3 Age, sex and ethnicity

Approximately one half of all type 1 diabetes cases are diagnosed before the age of 15 years (Karvonen et al., 2003). Before this time, there is an increase in incidence with age up to puberty with peak incidence usually occurring between 10 and 14 years (DIAMOND, 2006, Slama, 2003). This peak has been shown to be slightly earlier in girls than boys (Campbell-Stokes and Taylor, 2005, Nystrom et al., 1992, Staines, 1996, Taplin et al., 2005). There is also evidence to suggest that the age of onset is becoming younger, with increases in the rates of type 1 diabetes among the under five years age group recently noted (Karvonen et al., 1999, Rosenbauer et al., 1999, Schoenle et al., 2001, Wilson et al., 2007).

Prior to 1995, a higher risk of type 1 diabetes was noted in males in Europe and populations of European origin, particularly after puberty. Populations with a female excess in incidence were mainly of Black or Asian origin (Karvonen et al., 1997a). More recent analyses have only revealed significant male excesses in some European countries (Finland, Bulgaria, Greece, Switzerland, England) and significant female excesses in parts of China, Israel and Australia, and in African-Americans in Chicago (DIAMOND, 2006). A small female excess was observed for children with type 1 diabetes aged under 15 years in New Zealand 1999-2000, but this finding was not statistically significant (Campbell-Stokes and Taylor, 2005). In the Canterbury region of New Zealand a male excess of cases was found in all age groups and became more marked with increasing age for the period 1990-1999 (Willis et al., 2002b).

It has been shown that type 1 diabetes incidence among Caucasoid populations is higher than Mongoloids and Blacks throughout the world (Karvonen et al., 1997a, Rewers et al., 1990). In New Zealand, a lower incidence of type 1 diabetes has been reported in Māori and Pacific children compared to children of European descent (Campbell-Stokes and Taylor, 2005, Smith,

1987, Wu et al., 2005) especially in Canterbury (Willis et al., 2002b). However, increasing numbers of Māori and Pacific children presenting with type 1 diabetes have been noted in Auckland (Willis et al., 2005), and this is likely to continue due to the projected increases in both Māori and Pacific children in New Zealand over the next decade (Ministry of Social Development, 2002, Wu et al., 2005).

4.4 Temporal trends

The risk of childhood type 1 diabetes has been increasing worldwide since the 1950s. The most recent data revealed an average annual increase in global incidence of 3.4 percent between 1995 and 1999 (DIAMOND, 2006). In New Zealand, incidence has doubled in the last 20-30 years (Campbell-Stokes and Taylor, 2005). The Canterbury region of New Zealand witnessed an 11.9 fold increase in incidence from 2.4 cases per 100,000 person-years in 1970, to 28.6 cases per 100,000 person-years in 2004 (Willis et al., 2005). As well as secular rises in incidence, a number of shorter term increases have also been documented. Unusual outbreaks or 'epidemics' of type 1 diabetes have been reported in a number of countries (World Health Organization and Diamond Project Group on Epidemics, 1992). For example, in Canterbury between 1990 and 1992, type 1 diabetes incidence was double the rate recorded for the preceding eight years (Brown, 1993, Scott et al., 1992). Furthermore, shorter-term temporal variations in incidence have also been documented. For example, peaks of type 1 diabetes cases diagnosed in autumn and winter months have long been recognised (Gamble and Taylor, 1969) and supported by studies in both the north (Lévy-Marchal et al., 1995) and south hemispheres (Santosa et al., 2001, Willis et al., 2002a). In addition, children with type 1 diabetes are more likely to be born in the summer months (Laron, 2002, Rothwell et al., 1996, Songini et al., 2001, Willis et al., 2002a). However some studies found no seasonality effects (McKinney and EURODIAB Seasonality of Birth Group, 2001) or contradictory results (Karvonen et al., 1998, Laron et al., 2005).

4.5 Geography

Arguably one of the most important determinants of type 1 diabetes risk, is where a child lives (Brown, 1993). The incidence of type 1 diabetes has been found to vary geographically by world region, different countries within world regions, and by areas within countries. For the period 1990-1999, the highest incidence was observed in Europe (particularly northern and western Europe), Canada, North America and Oceania. The incidence of type 1 diabetes was generally

very low in Asia. For example, 70 percent of Asian populations observed an incidence of less than one case per 100,000 population per year. African populations for which data were available, recorded low to intermediate incidence rates (DIAMOND, 2006).

Considerable variation in the incidence of type 1 diabetes has also been reported between countries around the world (Karvonen et al., 2000). The most recent data available (1990-1999) at the national-level, revealed the lowest incidence of 0.5 cases per 100,000 per year in the Dominican Republic, and the highest incidence of 40.9 cases per 100,000 per year in Finland (Figure 4.1). Other countries with very high incidence included Sweden with 30.0, Kuwait with 22.3 and Northern Ireland with 21.3 cases per 100,000 per year. All of the countries with the top ten highest incidence rates were in Europe, with the exception of Kuwait in the Middle East (DIAMOND, 2006).

Large geographical variations have also been found within countries. For example, there was a 45-fold difference between the lowest incidence (0.1 cases per 100,000 per year) and the highest incidence (4.5 cases per 10,000 per year) rates of type 1 diabetes recorded in China during 1990-1999 (DIAMOND, 2006). A recent study in New Zealand found that children living in the South Island had a 1.5 fold higher incidence of type 1 diabetes than children living in the North Island during 1999-2000. The majority of North Island regions had an incidence of less than 20 cases per 100,000 population at risk, compared to the majority of South Island regions which had over 20 cases per 100,000. On the South Island, the region of Otago had the highest incidence of around 30 cases per 100,000 and the West Coast region had the lowest incidence since no new cases were diagnosed in this area over the study period. The Canterbury region had an incidence of around 23 cases per 100,000 population at risk (Campbell-Stokes and Taylor, 2005). However, it should be noted that these incidence rates were based on only two years of data, and that incidence rates of type 1 diabetes can vary considerably from year to year (Willis et al., 2005).

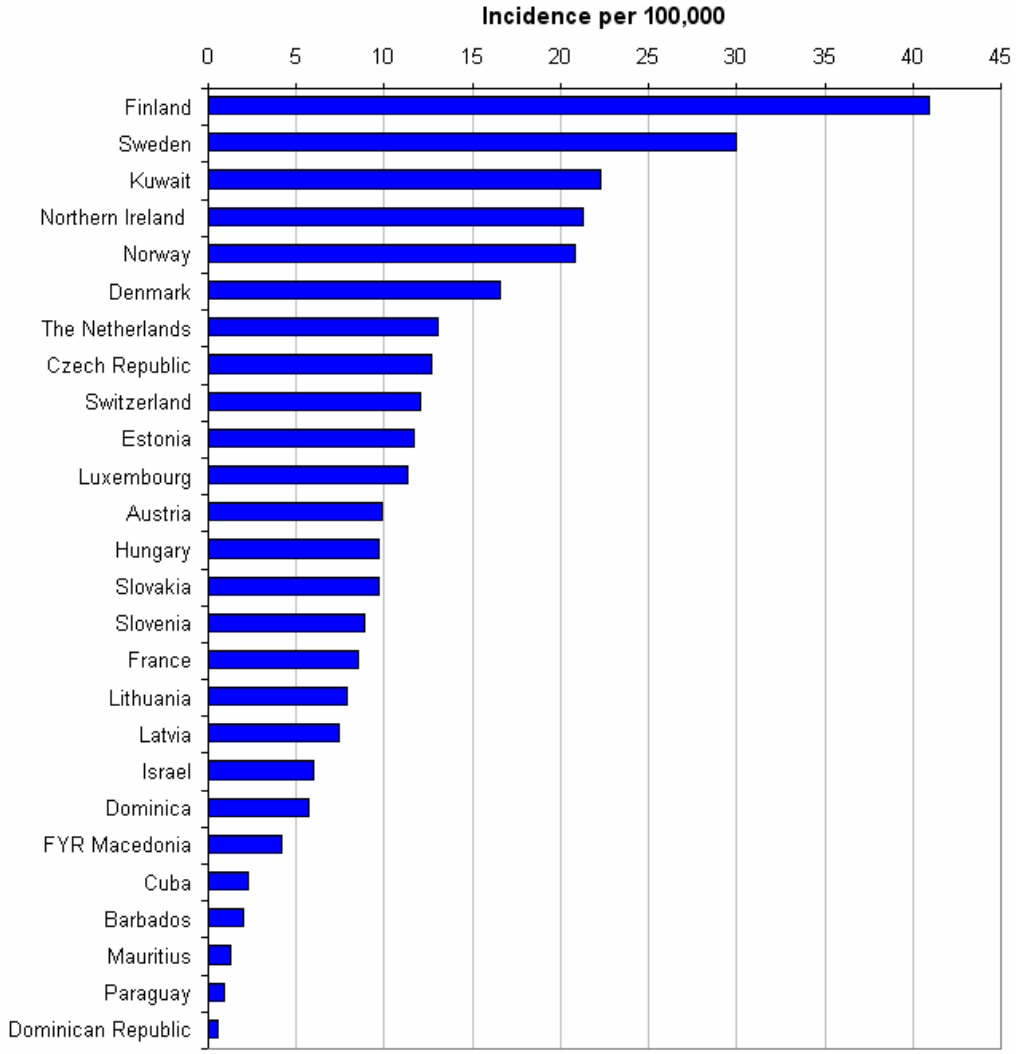

Figure 4.1: Type 1 diabetes incidence per 100,000/year by country 1990-1999
(adapted from: DIAMOND, 2006: 859)

4.6 Aetiology of type 1 diabetes

While geographical and temporal variations in type 1 diabetes incidence are well documented, the reasons for these differences and the aetiology of the disease in general, remain unclear. The current aetiological model for the development of type 1 diabetes contends that everyone is born with a degree of susceptibility to develop this disease, and that in susceptible populations, exposure to one or more environmental triggers is necessary to initiate beta cell destruction (Eisenbarth, 1986). This section outlines the main genetic and environmental risk factors that have been associated with type 1 diabetes development.

4.6.1 Genetic susceptibility

The importance of genetics in type 1 diabetes causation has been recognised for over 2000 years (Simpson, 1976). Family studies have revealed that direct relatives of type 1 diabetes patients have a greater risk of developing the disease than the general population. However, the risk is not the same for all first degree family members (Karvonen et al., 2001). For example, around two to three percent of children whose mother has type 1 diabetes are likely to develop the disease, compared to five to six percent of children whose father has type 1 diabetes. The risk rises to almost 30 percent if both parents have diabetes. For siblings of someone with type 1 diabetes the risk is around eight percent (EURODIAB, 1998, Lamb, 2006, Ziegler et al., 1999). Twin studies also emphasise the importance of genetic inheritance. In identical twins where one twin has type 1 diabetes, around 30-50 percent of the unaffected twins will eventually develop the disease (Genuth, 2006, Hyttinen et al., 2003).

Identifying the specific genes associated with increased diabetes risk is problematic due to the small number of cases where another immediate family member also has the disease. However, several important observations have been made (Achenbach et al., 2005). Genetic susceptibility has been found to be associated with Human Leucocyte Antigens (HLA) which are proteins on cell surfaces that induce the formation of antibodies to fight foreign cells after they enter the body (Brown, 1993). The types of HLA which individuals have are determined by a group of genes located on chromosome six (Karvonen et al., 2001). Some HLA sub-types may confer genetic susceptibility to type 1 diabetes by aiding recognition and presentation to the immune system of diabetogenic antigens (Brown, 1993). HLA class II molecules DR3 and DR4 are strongly associated with type 1 diabetes. More than 90 percent of Caucasoids with type 1 diabetes express one or both of these molecules, compared to 50-60 percent in the general population (Lamb, 2006). In addition, some protective HLA sub-types (e.g. genotypes containing the HLA DQ6 haplotype) have also been identified (Achenbach et al., 2005).

The worldwide variation in the incidence of type 1 diabetes is thought to reflect the global distribution of ethnic populations, and therefore to highlight the importance of differential genetic susceptibility. However, substantial differences in incidence have been reported in Caucasian populations living in relatively close proximity and among those who are genetically similar. For example, the incidence of type 1 diabetes in Finland is three and a half times higher than in Estonia (Karvonen et al., 2000). Furthermore, studies of migrating populations have found a convergence of incidence rates for immigrants with those of the host population

(Bodansky et al., 1992, Feltbower et al., 2002). For example, South Asian children living in Leicestershire in the UK had similar type 1 diabetes incidence rates to children from the local area. The high incidence rate for these children of around 20 cases per 100,000 population was in stark contrast to the low incidence of the disease reported in most Asian countries (DIAMOND, 2006, Raymond et al., 2001). Such findings suggest that environmental factors prevalent at the destination area are involved in causing type 1 diabetes.

4.6.2 Environmental risk factors

It has been estimated that between 60 and 95 percent of childhood diabetes worldwide, is triggered by environmental causes (Diabetes Epidemiology Research International Group, 1987). Unfortunately our knowledge regarding the effects of various environmental factors on this condition remains limited (Karvonen et al., 2000). Identification and understanding of the environmental causes of this disease is imperative since environments are potentially modifiable and may aid in diabetes prevention strategies (Slama, 2003). A number of possible environmental causes have been discussed in the literature, the most putative are reviewed here; diet, weight gain, chemical toxins and infections.

4.6.2.1 Diet

A variety of nutritional risk factors have been implicated in the aetiology of type 1 diabetes (Virtanen and Knip, 2003). In particular, infant diet has received considerable attention and focuses on the potential protective effects of breast milk and the possible harmful effects of cow milk. Increasing interest is also being given to the possible protective effects of vitamin D supplementation (Mathieu and Badenhoop, 2005).

Recent evidence suggests a role for vitamin D in the pathogenesis of type 1 diabetes (Mathieu and Badenhoop, 2005). Animal models have shown that vitamin D supplementation prevents autoimmune diabetes in non-obese diabetic (NOD) mice (Mathieu et al., 1994). Similar effects have been noted in human studies. For example, a large multi-centre trial in Europe revealed a protective effect (odds ratio = 0.67, 95 percent confidence intervals = 0.53-0.86) of vitamin D supplementation in infancy (EURODIAB Substudy 2 Study Group, 1999). Furthermore, in a birth cohort study in northern Finland, children who regularly took the recommended dose of vitamin D (2000 IU daily) had a significantly reduced risk of developing type 1 diabetes (relative risk = 0.22, 95 percent confidence intervals = 0.05-0.89) compared with children who regularly

received less than the recommended amount (Hypponen et al., 2001). In addition, maternal intake of vitamin D during pregnancy has been shown to protect against islet autoimmunity in offspring (Brekke and Ludvigsson, 2007). However, genetic data on the association between type 1 diabetes and vitamin D remain controversial, warranting further research in this area (Gillespie, 2006).

The possible links between type 1 diabetes and breast-feeding have received considerable attention. Borch-Johnsen et al (1984) were the first to report an inverse correlation between breast-feeding frequency and duration, and type 1 diabetes development in Copenhagen. Children with diabetes were breast fed for shorter periods of time than their siblings without the disease and the background population. Since this work, several studies in various countries have found evidence to support this association (e.g. Blom et al., 1989, EURODIAB, 2002, Hummel et al., 2007, Kimpimäki et al., 2001, McKinney et al., 1999b, Patterson et al., 1994, Siemiatycki et al., 1989, Virtanen et al., 1991) with some exceptions (e.g. Couper et al., 1999, Hummel et al., 2000, Ziegler et al., 2003). These studies are reviewed in Akerblom and Knip (1998) and Virtanen and Knip (2003). A number of mechanisms have been proposed for how breast milk might confer protection against type 1 diabetes in children. Breast milk is known to protect new born infants against infections through secretory immunoglobulin A antibodies, and to enhance the infant's own immune responses (Borch-Johnsen et al., 1984). Furthermore, increased beta cell proliferation has been associated with breast fed children compared to those fed on formula milk (Juto, 1985), and prolonged breast-feeding may delay exposure to foreign food antigens (Virtanen and Knip, 2003).

Proteins found in cow's milk have also been linked with the development of diabetes in humans (Borch-Johnsen et al., 1984, Gerstein, 1994, Norris and Scott, 1996), diabetes prone bio-breeding (BB) rats (Elliot and Martin, 1984) and NOD mice (Karges et al., 1997). In humans, increased numbers of antibodies toward a series of cow milk proteins have been detected in children with newly diagnosed type 1 diabetes (Glerum and Robinson, 1989, Savilahti et al., 1988, Virtanen et al., 1994b). For example, recent debate has questioned whether the specific A1 protein in cow's milk could facilitate the immunological process which leads to type 1 diabetes (Allison and Clarke, 2006, Elliott et al., 1999, Truswell, 2005, Woodford, 2006). Furthermore, early introduction of cow's milk has been associated with an increased risk of type 1 diabetes (Gerstein, 1994, Hypponen et al., 1999, Scott et al., 1996). Preliminary findings from a randomised trial in Finland suggest that replacement of dietary cow's milk proteins with hydrolysed casein formula at 6-8 months protects genetically at risk children from the emergence

of signs of beta cell autoimmunity over the first two years of life (Akerblom et al., 1999). Several theories have been proposed to explain the diabetogenicity of cow's milk (Virtanen and Knip, 2003). In the developed world, cow's milk contains the first foreign proteins that children ingest and is hence an early test of a child's developing immune system (Akerblom et al., 2002). It is possible that cow's milk proteins cause direct harm to islet beta cells or that bovine insulin in milk operates as an antigen against beta cells (Karvonen et al., 2003). However, contradictory results have been found with some studies finding no significant relationship between the development of islet autoimmunity and the introduction of cow's milk protein in the infant diet (e.g. Couper et al., 1999). Possible explanations for the mixed results include differences in study design especially in measuring the timing and duration of exposure, differential inclusion of confounding variables and genetic variation in cow's milk proteins (MacFarlane and Scott, 2003).

4.6.2.2 Weight gain

Cow's milk protein may also be a marker for some other unmeasured factor. For example, weight gain is likely to be greater in infants fed cow's milk protein-based formula compared to those who are primarily fed breast milk (Akerblom and Knip, 1998). Evidence to substantiate the importance of weight gain in infancy for type 1 diabetes pathogenesis has been shown in a number of studies (Bruining, 2000, EURODIAB, 2002, Hypponen et al., 1999, Johansson et al., 1994). In addition some studies have also found that weight gain in childhood is predictive of type 1 diabetes occurrence. A study in Finland found that obesity after three years of age was associated with a more than two-fold risk of developing type 1 diabetes (Hypponen et al., 2000). However, some childhood studies have found weight gain to be unrelated to type 1 diabetes risk (Blom et al., 1992, Bruining, 2000).

One hypothesis put forward to explain how weight gain could cause type 1 diabetes, is the 'accelerator hypothesis'. This was proposed by Wilkin (2001) who contends that type 1 and type 2 diabetes are the same disease, distinguishable only by the rate of beta cell loss and the accelerators that cause this. Wilkin identifies three accelerators: constitution, insulin resistance and autoimmunity. Insulin resistance typically results from weight gain, is thought to contribute to beta cell death and is central to the link between the two types of diabetes. Autoimmunity is only present in individuals with a genetic predisposition to beta cell autoimmunity and where this accelerator is absent, beta cell death is slower and progression is to 'type 2 diabetes'. The accelerator hypothesis is consistent with secular increases in both childhood obesity (Chinn et

al., 1998, Troiano et al., 1995) and type 1 diabetes (DIAMOND, 2006) that have occurred over the last 30 years in the developed world. In addition, the decreasing age at diagnosis noted in some populations (e.g. Dabelea et al., 2006, Pundziute-Lycka et al., 2002) may reflect the heavier weight of children leading to earlier diabetes presentation (Daneman, 2005). Studies which have sought to test this hypothesis have found both supportive (Betts et al., 2005, Fourlanos et al., 2004, Kibirige et al., 2003) and contrary results (O Connell et al., 2007). However, the increase in weight could be a maladaptation to pre-diabetes accelerated by another cause (Kibirige et al., 2003). Furthermore, the pathophysiology and natural histories of type 1 and type 2 diabetes are quite distinct, making it problematic to accept that they are the same disease with the same aetiology (Daneman, 2005).

4.6.2.3 Toxins

Ecological, case-control and animal studies have also implicated the ingestion of N-nitroso compounds in the aetiology of type 1 diabetes. The most important exogenous source of these compounds is food. Nitrate and nitrite are found in food as naturally occurring compounds, but are also added in the processing of meat products to delay decay and to improve colour and taste (Virtanen and Knip, 2003). These compounds are also found in fertilisers used for farming and can build up in water sources. Therefore this risk factor is more likely to be important in rural/agricultural areas where contamination of wells by surface water occurs. It is thought that N-nitroso can directly damage pancreatic beta cells or initiate an autoimmune response against these cells (MacFarlane and Scott, 2003).

Previous studies investigating the link between type 1 diabetes and N-nitroso display mixed results. In Colorado, USA, a low positive correlation was reported between the nitrate content of drinking water and the incidence of childhood type 1 diabetes at the county-level (Kostraba et al., 1992). A study from England found a significant excess of childhood diabetes cases living in areas with the highest mean nitrate levels (Parslow et al., 1997). A number of case-control studies have supported these findings, reporting higher intake of nitrates and nitrites in children with diabetes compared to other children (Dahlquist et al., 1990, Virtanen et al., 1994a). However, Canadian (Siemiatycki et al., 1989) and Australian (Verge et al., 1994) studies did not find any significant differences in N-nitroso intake between case and control children. In addition, ecological studies in Sardinia (Casu et al., 2000) and the Netherlands (Van Maanen et al., 1999) have been unable to confirm the positive association between childhood diabetes and nitrates in drinking water. As a result of these mixed findings, further investigation is warranted.

Of particular interest is the dose needed to instigate type 1 diabetes onset, the importance of dose timing, and the potential cross-reaction with other risk factors (Akerblom et al., 2002).

4.6.2.4 Infections

Infections have long been implicated in the aetiology of type 1 diabetes (Harris, 1899, Kremer, 1947, Maugh, 1975), but proving a causal role between specific infections and type 1 diabetes onset has proven difficult (Lammi et al., 2005). This is perhaps because the links between infection and autoimmunity are likely to be more complex than were originally assumed (Filippi and von Herrath, 2005). As well as acting as a trigger for type 1 diabetes, it has also been argued that infections can protect against, or have no effect, on the development of the disease depending on a person's immune status (Bach, 2005a, Filippi and von Herrath, 2005).

A triggering role for infections

A number of viruses have been associated with triggering the development of type 1 diabetes (Filippi and von Herrath, 2005). Evidence for this association has been provided in studies of recently diagnosed patients with serological evidence of viral infection, and in several case reports of viral infections preceding type 1 diabetes onset (Yoon and Jun, 2003). Supporting observations have been found in epidemiological studies regarding seasonal (Lévy-Marchal et al., 1995, Willis et al., 2002a) and geographical variations in incidence (e.g. Cherubini et al., 1999, Rytkonen et al., 2001, Samuelsson and Lofman, 2004), including space-time clustering of the disease. For example, previous studies in Sweden (Samuelsson and Carstensen, 2003, Samuelsson et al., 1994), Chile (Santosa et al., 2001) and the United Kingdom (Law et al., 1997, McNally et al., 2006a) have all found significant clusters of type 1 diabetes by address at diagnosis. Other studies have found clustering of the disease by place of birth (Dahlquist and Kallen, 1996) and primary school attended (Bodington et al., 1995). These findings are consistent with an infectious aetiology. Animal models have also played an important role and have identified a number of diabetogenic viruses in mice and rats (Jun and Yoon, 2001). Specific viruses implicated in the development of this disease include the rubella virus, mumps virus, rotavirus and enteroviruses such as the coxsackievirus B (van der Werf et al., 2007).

There is substantial evidence implicating rubella as a causative factor in human diabetes (Forrest et al., 1967, Ginsberg-Fellner et al., 1980, Hyöty and Taylor, 2002, Lammi et al., 2005). Up to 20 percent of patients with congenital rubella have been found to later develop diabetes (Shaver

and al, 1985). However, since the widespread use of its vaccine, this virus is unlikely to result in many type 1 diabetes cases in the future (Hyöty and Taylor, 2002). Similarly, the mumps virus, which was one of the first to be associated with type 1 diabetes in humans (Harris, 1899, Hyoty and al., 1988, Kremer, 1947), is unlikely to play a major role in future diabetes causation (Hyöty and Taylor, 2002). Rotavirus studies have given mixed results. One small study has suggested that this virus might induce type 1 diabetes in genetically susceptible children (Honeyman et al., 2000), while a larger, more recent study found no association (Blomqvist et al., 2002). This later study has however, been criticised for defining and measuring rotavirus infection incorrectly (Honeyman, 2005).

While it has not been confirmed that enteroviruses are aetiologic agents of type 1 diabetes (Filippi and von Herrath, 2005), a number of studies have shown that these viruses accompany or precede onset of this disease in children (Hyöty and Taylor, 2002). Coxsackievirus B (CVB) is most often implicated and several epidemiological studies have reported high frequencies of anti-coxsackie B antibodies in children with newly diagnosed type 1 diabetes (e.g. Gamble et al., 1973). For example, significant increases in incidence of type 1 diabetes were observed in Jefferson County, Alabama after an epidemic of CVB 5 which occurred in 1983 (Wagenknecht et al., 1991). In addition, isolates of CVB 4 and 5 have been shown to induce diabetes in susceptible mouse strains (Champsaur et al., 1982, Yoon and Jun, 2003). However, a recent study has shown a significant decrease in the prevalence of maternal enterovirus antibody levels over the last 20 years, and also that these viruses are less prevalent in high incidence countries (e.g. Finland) when compared to low or intermediate incidence countries (e.g. Estonia and Karelia) (Viskari et al., 2005). This finding suggests that coxsackieviruses are not linked to temporal increases in type 1 diabetes incidence, and also that other risk factors must be involved in areas of high type 1 diabetes prevalence.

As well as uncertainty regarding the specific viruses involved in triggering type 1 diabetes, the mechanisms by which viruses can induce this disease are complex and still under debate. They can be divided into two main categories; direct cytolysis (cell destruction) of virus infected cells with no involvement of the immune system, and the viral induction of autoimmune processes (Lammi et al., 2005). Several viruses are able to directly infect beta cells in the pancreas, and some infections can lead to beta cell destruction (Filippi and von Herrath, 2005). For example encephalomyocarditis has been shown to induce diabetes in mice without involving the immune system (Jun and Yoon, 2001, Yoon et al., 1985). In humans, CVB strains have been shown to cause functional impairment and beta cell death characterised by shrinking of the cell nucleus

(Roivanen et al., 2000). However, CVB 4 has been associated with tissue damage in mice which can result in the release of autoantigens from the beta cells, constituting an autoimmune response (Horwitz et al., 2002).

Several explanations regarding the autoimmune mechanisms involved in the pathogenesis of type 1 have emerged (Lammi et al., 2005). One hypothesis is molecular mimicry, where similar structures are shared between molecules from dissimilar genes. For example, if during a viral infection, the infecting organism shares cross-reactive epitopes for B or T cells with the host, then the response to the infecting agent will also attack the host, causing autoimmune disease (Oldstone, 2005). A significant homology has been found between the CVB 4 protein and the glutamic acid decarboxylase (GAD) sequence in the beta cell islets (Atkinson et al., 1994). However, T cells have not been found to cross-react to GAD *in vitro* so this mechanism remains unproven (Lammi et al., 2005). Another hypothesis formed to explain autoimmune mechanisms involves the bystander activation of T lymphocytes (Filippi and von Herrath, 2005). Viral infection could result in inflammatory effects which induce tissue damage and release sequestered islet antigens resulting in the re-stimulation of resting autoreactive T cells (Horwitz et al., 1998). Infections have also been implicated in type 1 diabetes development through their role in disturbing the balance between two types of T helper cells (Th). The hypothesis postulates that destruction of beta cells occurs when this balance is shifted towards diabetogenic Th1 cells, and away from protective Th2 cells (Singh et al., 1998). However, it has recently been suggested that this argument is too simplistic, and that Th1 cells are not the sole instigators of type 1 diabetes and that Th2 cells are more harmful than previously believed (Azar et al., 1999). The data currently available assessing the effects of T cell balance is still sparse and further work is necessary (Lammi et al., 2005).

Despite intensive research, conclusions regarding specific infections and mechanisms of type 1 diabetes development have not yet been made. One of the most important problems which hinder studies on this topic remains the issue of timing. The symptoms of type 1 diabetes occur at the end of a pre-clinical period during which the body's immune system destroys the insulin producing beta cells in the pancreas. This pre-clinical period can last months or years (Honeyman, 2005). Thus the factors responsible for initiating the pathogenic process may only be identifiable years before the onset of symptoms, and therefore may no longer be detectable at the time of diagnosis (Lammi et al., 2005). In addition, infections are transient in nature and the detection of antiviral antibodies in type 1 diabetes patients has always been elusive (Bach, 2005a).

A protective role for infections

In addition, there is increasing evidence to support a protective role for infections in the development of type 1 diabetes (Bach, 2001, Kolb and Elliot, 1994). Research investigating the role of the 'hygiene hypothesis' in the pathogenesis of this disease is beginning to emerge. The theory behind the hygiene hypothesis was first expressed by Strachan in 1989 in relation to hay fever (Strachan, 1989) but has since been extended to include various autoimmune diseases such as type 1 diabetes (Bach, 2005a). This hypothesis states that improved hygiene in Western countries has resulted in a decrease in common childhood infections which could explain the dramatic increase in the incidence of autoimmune diseases. It is thought that early exposure to common infections is beneficial for the immune system, reducing the chances of autoimmune responses in later life (Lammi et al., 2005).

Specific biological mechanisms by which infections can protect against type 1 diabetes development are, however, unclear and are likely to be multifactorial. One hypothesis is that antigenic competition occurs: strong immune responses that are elicited by infectious agents compete with immune responses against weaker antigens, such as autoantigens. Thus, the immune system is too busy fighting an infection to attack the body (Bach, 2005a). Another possibility is that infectious agents stimulate the production of regulatory cells whose effects extend beyond responses to the invading microbe. It is thought that these cells help to dampen autoimmune responses (Bach, 2002). In addition, Toll-like receptors are thought to play a central role in the stimulation of autoimmune responses. For example, the administration of various Toll-like receptor agonists in young NOD mice have been shown to prevent diabetes onset (Bach, 2005a).

Evidence to support a protective role for infection is supplied by animal and epidemiological studies. It has been consistently observed that animals bred in pathogen free environments have higher rates of autoimmune disease than those bred under normal conditions (e.g. Like et al., 1991, Ohsugi and Kurosawa, 1994). For example, the use of caesarean section and isolated living conditions has been found to increase the incidence of diabetes in NOD mice (from 50 percent) by two to three fold in a single generation (2001). Moreover, diabetes risk has been found to be lower in animals exposed to infections in early life (Bach, 2002). The inoculation of diabetes-prone BB rats at 30 days of age with a lymphotropic variant of the lymphocytic choriomeningitis virus has been shown to significantly reduce the incidence of diabetes (Schwimmbeck et al., 1988). In addition, diabetes is prevented in NOD mice by infecting the

young mice with filariae (Imai et al., 2001) and various other viral and bacterial infections (Bach, 2002).

In humans, infections in early life have also been associated with a reduction in type 1 diabetes risk. A study of 58 children with type 1 diabetes and 172 age- and sex-matched controls in Southampton in England found that infections had a protective effect on type 1 diabetes incidence within the first year of life (Gibbon et al., 1997). In a larger case-control study in Lithuania (124 type 1 diabetes patients and 372 controls), children were also less likely to develop diabetes if they experienced one or more infections in the first six months after birth (Pundziute-Lycka et al., 2000). A case-control study in Yorkshire in the UK investigated the link between childhood type 1 diabetes and early social mixing as measured by attendance at day-care facilities. Pre-school day-care attendance is known to be associated with an increased burden of infectious disease (Osterholm, 1994), and therefore should be associated with a decreased risk for type 1 diabetes. The Yorkshire study supported this theory; attendance at day-care centres for infants below one year of age showed a significant inverse association with childhood diabetes (McKinney et al., 2000). Pre-school day-care attendance was also found to be inversely associated with childhood diabetes in data from seven European population-based registers. However, this study also noted that infections early in the child's life were associated with an increased risk of diabetes after adjustment for confounding variables (EURODIAB, 2000). A recent review found that the majority of case-control studies showed a statistically significant protective effect of day-care attendance, but a meta-analysis revealed too much heterogeneity to accept the overall synthesis of results (Kaila and Taback, 2001).

Early childhood infections are also contracted from siblings. Higher birth order was associated with a decreased risk in type 1 diabetes in a national case-control study in the UK (Wadsworth et al., 1997). This finding has been supported by a prospective population based family study in Oxford, where the risk of type 1 diabetes was significantly higher for first born children after controlling for maternal age (Bingley et al., 2000). In Northern Ireland, children of first pregnancies were also at increased risk of developing this disease (Patterson et al., 1994). This result has been confirmed by more recent work, but only in children diagnosed under the age of 5 years (Cardwell et al., 2005). Outside the UK, lower incidence of type 1 diabetes has also been associated with increasing birth order in Western Australia (Haynes et al., 2007a). However, no association was found with maternal parity in Denmark (Bache et al., 1999).

Measures of household overcrowding have also been used as a proxy for infectious exposure. Overcrowded conditions have been strongly related to the spread of infectious diseases in children (Smith et al., 1990). Work in Northern Ireland found that children living in areas with medium and high levels of household overcrowding (measured as the number of people per room) had a significantly reduced risk of childhood diabetes when compared to those living in areas of low household overcrowding (Patterson et al., 1996). Similarly, children living in areas with many overcrowded houses in Yorkshire in England had significantly lower rate ratios of type 1 diabetes for 1978-1990 (Staines et al., 1997). However, two case-control studies using data from the 1980s have reported an increased risk in children living in crowded households in early life (Lawler-Heavner et al., 1991, Siemiatycki et al., 1989).

Conditions related to a child's area of residence have also been used as a measure of circulating infections. For example it is known that remote rural areas with low population densities and migration rates are less likely to sustain an extensive range of endemic infections (Rhodes and Anderson, 1996). As a result, many studies have tested the relationship between urban/rural status or area-level population density and type 1 diabetes incidence. Studies carried out in the UK have generally shown lower incidence of type 1 diabetes in urban, compared to rural, areas. For example, type 1 diabetes standardised incidence ratios were 20 percent lower in urban, compared with rural, postcode sectors in Scotland between 1977 and 1983 (Patterson and Waugh, 1992). A study in Yorkshire (1978-1990) also found deficits of childhood diabetes in urban areas (McKinney et al., 1996) and later work in this region supported these early findings (Feltbower et al., 2005, Parslow et al., 2001, Staines et al., 1997). Furthermore, in Northern Ireland during the period 1989-1994, there was a reduced risk of diabetes for children living in medium and high population density areas relative to those living in areas of low population density (Patterson et al., 1996). More recent studies in Northern Ireland have revealed compatible results (Cardwell et al., 2006, Cardwell et al., 2007). Studies in Finland generally support the UK findings (Karvonen et al., 1997b, Rytkonen et al., 2003). However, other European studies have revealed contradictory results. Childhood diabetes incidence was found to increase with the degree of urbanisation in central-eastern and south-western Italy for the period 1990-1995 (Cherubini et al., 1999). Moreover, incidence of this diseases was higher in metropolitan areas compared to urban, rural, and semi-rural areas in Greece (Dacou-Voutetakis et al., 1995). However, nationwide studies in Switzerland (Schoenle et al., 2001) and Austria (Schober et al., 2003) found no significant association between childhood diabetes and urban/rural status or population density. To date, no studies have investigated this issue in New Zealand, but in Western Australia, type 1 diabetes incidence was higher in urban compared to

non-urban areas (Haynes et al., 2007a, Haynes et al., 2006). Thus the relationship between type 1 diabetes incidence and the level of urbanity/rurality (or population density) seems to be country specific. These inconsistencies could be due to different types of urban/rural landscapes in different countries with potentially differing population densities and migration rates, inconsistencies between how these features are measured, or differences in other unmeasured risk factors prevalent in these areas.

It is argued that children from higher social classes are likely to have reduced or delayed exposure to infectious agents due to increased hygiene levels (Tedeschi and Airaghi, 2006). Consequently, several studies have also investigated the relationship between type 1 diabetes and deprivation. At the national-level, affluence as measured by gross national product was positively correlated (correlation coefficients of 0.58 and 0.53 respectively) with the incidence of type 1 diabetes in several European countries (Patterson et al., 2001, Tedeschi and Airaghi, 2006). Comparable results have also been found for smaller geographical areas. For example in Western Australia, significantly higher incidence of type 1 diabetes was found in children living in the most affluent census collection districts of the region (Haynes et al., 2006). Results from ecological studies in England (Feltbower et al., 2005, Parslow et al., 2001), Northern Ireland (Cardwell et al., 2006, Patterson et al., 1996) and Scotland (Patterson and Waugh, 1992), also revealed compatible results. Reduced type 1 diabetes incidence was observed in the most deprived areas in each study. However, some studies do not support these findings. For example, a significant positive trend was noted between material deprivation (Townsend score) and childhood diabetes incidence in northern England 1977-86 (Crow et al., 1991). A more recent study in Germany also revealed higher deprivation scores to be associated with an increased risk of this disease (du Prel et al., 2007). In addition, a number of other studies (e.g. Baumer et al., 1998, Cox, 2007, Evans et al., 2000) have found no association between deprivation and type 1 diabetes incidence. The conflicting findings may result from incomparable study designs, for example the use of different deprivation measures or geographical scales of analysis, or due to differences in gene-environment interactions (du Prel et al., 2007). Nonetheless, the majority of the studies have shown childhood diabetes to be higher in more affluent areas, which is suggestive of a protective role for higher levels of infections in more deprived settings.

Population mixing

A relatively new approach to measuring area-level infectious exposure involves the use of population mixing measures. Population mixing has been used as a proxy for the number and

range of circulating infections in studies exploring the aetiology of acute lymphoblastic leukaemia for some time (Chapter 3), but has only recently been employed in type 1 diabetes research. To date, there are only three area-level studies on population mixing and type 1 diabetes; two of which consider small area variation in type 1 diabetes incidence in Yorkshire in the UK. In the first Yorkshire study, electoral wards with low levels of population mixing (bottom decile) in 1991 were associated with significantly higher incidence rates of childhood diabetes for 1986-1994. Children over four years of age were at greatest risk (Parslow et al., 2001). The second Yorkshire study used the same measures of population mixing and areas of analysis, but incorporated more recent diabetes data into the models (1986-1998) and used different statistical methods of analysis. The results supported the findings of the previous study; significantly higher childhood diabetes incidence was found in areas with the lowest population mixing levels (Feltbower et al., 2005). Comparable results have also been found by research in Tayside, Scotland, which investigated the association between childhood type 1 diabetes and a number of different measures of population mixing. Areas with a higher percentage of child in-migrants had significantly lower rates of type 1 diabetes, and areas with both low child in-migration *and* low child migrant diversity had the highest incidence of type 1 diabetes (Cox, 2007).

While these studies have tested a theory regarding what makes children more susceptible to type 1 diabetes, they have not tested what the final trigger to overt disease may be. In a discussion of their results, Parslow et al (2001) mention that a 'late exposure' to infections may provide the necessary 'hit' to the weakened immune system to cause type 1 diabetes (p.537). However, the authors did not test this theory. This hypothesis is very similar to that proposed in the aetiology of childhood leukaemia, with the majority of studies supporting the concept (Chapter 3). Such studies suggest that the timing of exposure to infections is critical, and may help reconcile the potentially contradictory findings regarding infections in the aetiology of type 1 diabetes.

4.7 Conclusion

In recent years, the incidence of childhood type 1 diabetes has been rising in New Zealand and in many other affluent countries. The possible causes of the disease and its increasing incidence have received considerable attention in the literature. However, many questions remain unanswered. This chapter has summarised the current knowledge regarding the aetiology of type 1 diabetes, highlighting the role for both genetic and environmental risk factors. Of particular importance to the thesis, is a growing body of evidence implicating infections in the aetiology of

this disease. Data support both a triggering and protective role for infections and further work is necessary to elucidate these associations.

A recent development in this area of research is the use of population mixing as a proxy measure for area-level infectious exposure. Studies on population mixing and childhood type 1 diabetes have consistently shown that diabetes incidence is reduced in areas of high population mixing. However, it is also possible that infections can trigger type 1 diabetes. Consequently, an increase from low to high population mixing levels could result in increased type 1 diabetes incidence, as shown in childhood leukaemia studies (Chapter 3). However, to date, no studies have tested this theory. Furthermore, the importance of population mixing in the aetiology of childhood type 1 diabetes needs to be examined in a wider variety of geographical settings. Only three studies have been carried out to date; all in the UK. New Zealand is characterised by high population mobility which varies spatially around the country and is thus an ideal location to explore these theories further. The next chapter discusses the data and methods employed in this research to examine the role of population mixing in childhood health in the New Zealand context.

Chapter 5 - Data and methods

5.1 Introduction

Using information from the literature on population mixing (Chapter 2), ALL (Chapter 3) and type 1 diabetes (Chapter 4), a number of quantitative methods were employed to address the aims of this thesis. This chapter begins with a description of the population mixing, ALL and type 1 diabetes data used in the research. The remainder of the chapter details the methods used in two sub-sections. First, the methods utilised to determine the basic epidemiology and geography of both diseases in the New Zealand setting are summarised. Second, the approaches used to statistically examine the relationship between population mixing and each disease are detailed. A number of population mixing measures are formulated to capture a range of population movements in New Zealand. Exploratory methods that are used to begin examining the associations between population mixing and ALL/type 1 diabetes are then described. Finally, the datasets and modelling strategy employed in the regression analyses of this research are detailed.

5.2 Data

5.2.1 Population mixing data

Since this research aimed to assess the effects of population mixing across the whole of New Zealand, accurate counts of population movements were required at a small spatial scale, for the whole country, at regular time intervals. The only data source that met those requirements was the national census. The census is the official count of population and dwellings in New Zealand, and provides a 'snapshot' of various aspects of society every five years (Statistics New Zealand, 2001). Responses to a number of questions in the census provide information on both internal and international in-migration, and the movements of overseas visitors around the country. This study utilised data collected on individuals' usual residence five years ago, including information on which of the 74 territorial authorities (TA) in New Zealand if they were internal movers, and which overseas country if they had emigrated to New Zealand within the past five years. Data were also obtained on the number of years individuals had resided at their current residence, and also on the number of overseas visitors present in New Zealand on census night. All of these data were supplied by Statistics New Zealand for the years 1981, 1986, 1991, 1996 and 2001 and were aggregated to the 2001 census area unit (CAU) boundaries. CAUs are the second smallest

unit of dissemination of census data in New Zealand and each CAU represents approximately 2,000 people (Statistics New Zealand, 2006a).

5.2.2 Acute lymphoblastic leukaemia data

Data on childhood acute lymphoblastic leukaemia registrations in New Zealand between 1980 and 2004 were extracted from the New Zealand Cancer Registry (NZCR) by the New Zealand Health Information Service (NZHIS). The NZCR is a population-based register of all primary malignant diseases diagnosed in New Zealand, excluding squamous cell and basal cell skin cancers. Data sources include laboratory test results, medical certificates of causes of death, coroners' findings, hospital discharge data on the National Minimum Dataset, and private hospital discharge returns (NZHIS, 2004). The NZCR was originally set up in 1948 and has been regarded as truly population based since 1974 (Cooke et al., 1988). The completeness of case ascertainment for childhood cancers was independently assessed using capture-recapture methods for the period 1990-1993. During this time, the NZCR identified 98.6 percent of the new childhood leukaemia cases diagnosed throughout the country (Dockerty et al., 1997). Moreover, since the Cancer Registry Regulations came into effect in 1994, laboratories are required by law to report any new diagnosis of cancer within New Zealand. As a result, data quality and completeness have significantly improved (NZHIS, 2004). For this study, data on all children aged 0-14 years at diagnosis of acute lymphoblastic leukaemia (ICD10 codes C91.00 and C91.01) and usually resident in New Zealand were obtained for the years 1980 to 2004 (inclusive). More specifically, information on each child's sex, ethnicity, date of registration, age at registration, and census area unit of residence at registration, were provided. A total of 781 new cases of ALL were registered during this period of which 456 were male, and 325 were female (Table 5.1).

Geocoding

Initially, the CAU of residence at diagnosis was provided by the NZHIS for every ALL case. However, the level of accuracy of NZHIS geocoding is thought to only be around 75 percent (Ministry of Health, 2005) which was deemed unacceptable for this research. As a result, the full addresses at diagnosis were sought. Due to confidentiality requirements, these data were extracted and geocoded in conjunction with Craig Wright, a Senior Advisor at the Public Health Intelligence unit of the Ministry of Health.

Unfortunately, the full addresses at ALL registration were only available from the Cancer Registry for the years 1980-1992 and 2002-2004. The full addresses for the period 1993-2001 were accidentally lost when the registry changed database systems and the new database storage system did not contain an address field. To overcome this problem, address data for cases diagnosed between 1993 and 2001 were sought from two additional sources: the Mortality Collection from the NZHIS, and the National Health Index (NHI). The three address files were merged by the NHI number and where there was a full address from the cancer registry, this address was used. If there was no cancer registry address and the person died within 3 years of the cancer registration, then the address data from the mortality registry was used. Where no cancer registry address was available and the person did not die, or did not die within 3 years of registration, then an NHI address was used for geocoding. Since there were two available NHI address files, the last updated columns were compared to the date given for the registration of the leukaemia case. The closest date to the registration indicated which NHI address to use for geocoding for each case (C Wright 2006, pers. comm., 5 November).

The resulting access file was batch geocoded. Interactive geocoding was carried out on the addresses which were unmatched by the batch geocoding, or matched at the suburb/locality or town/city accuracy level.

Table 5.1: Disease data summary

1980-2004	Acute lymphoblastic leukaemia (New Zealand)	Type 1 diabetes (Canterbury)
Male cases	456	164
Female cases	325	173
0-4 years	436	78
5-9 years	207	101
10-14 years	138	158
Total cases	**781**	**337**

5.2.3 Type 1 diabetes data

Unlike cancer registrations and mortality records, cases of diabetes in New Zealand are not available at the national-level, and thus access to data for the whole country was not possible. As a result, a number of different sources of diabetes data were considered for this research. The largest data collection system currently in place is the 'Get Checked' programme which started in 2000. This initiative requires participating general practices to forward data on their diabetes

patients to a database administered by their local Primary Health Organisation (PHO). At present, there are 77 individual PHOs across New Zealand whose diabetes databases are subject to varying levels of completeness (S Dawson, 2005, pers. comm., 24 June), due to lower than anticipated levels of patient recruitment (Berkeley and Lunt, 2006). A further limitation of this data source is that it does not include children with type 1 diabetes who are admitted to hospital without visiting a general practitioner (S Dawson, 2005, pers. comm., 24 June).

National hospital admissions data were also considered. However, these data do not include patients treated at outpatient clinics, where large numbers of type 1 diabetes patients are managed (P Moore 2005, pers. comm., 19 August). Moreover, these data only include the patients' current address from their National Health Index (NHI) records, and inconsistent use of coding systems prior to 1995 has resulted in an underestimation of type 1 diabetes discharges by up to 20 percent for these years (C Lewis 2005, pers. comm., 16 August). Since national-level data were not of sufficient quality, access to two regional registers in the South Island was sought: SouthLink Health in Dunedin and the Canterbury Diabetes Register. Unfortunately, the SouthLink Health database did not contain each patient's address at diagnosis, only their current (NHI) address which is updated every three months and therefore was not appropriate for this research.

The Canterbury Diabetes Register, however, does contain each patient's original address at diagnosis of type 1 diabetes. Moreover, this register is thought to have 100 percent ascertainment of cases since 1970. The register covers five district council areas (Christchurch City, Banks Peninsula, Waimakariri, Hurunui and Selwyn) located in the Canterbury region of the South Island of New Zealand (Figure 5.1). The total population of the study area in 2001 was 408,606 or 10.7 percent of the total New Zealand population. This region has been identified by previous research as having relatively high rates of childhood type 1 diabetes (Campbell-Stokes and Taylor, 2005, DIAMOND, 2006, Willis et al., 2002b).

Prospective ascertainment of type 1 diabetes patients began in this region in 1982 using notifications from the patient's usual diabetes physician or paediatrician. All new children and adolescents (0-19 years) who develop the disease within this region are treated at the Christchurch regional hospital. Patients diagnosed prior to 1982 were ascertained retrospectively from hospital records at in-patient and out-patient clinics and were verified by case note reviews. Secondary data sources included membership of local lay diabetes societies, GP surveys, a

survey of insulin users at retail pharmacies and through direct communication with patients and families (Brown and Scott, 1988, Willis et al., 2002b).

For the purposes of this research, individual-level data for all children under the age of 15 who were diagnosed with type 1 diabetes between 1980 and 2004 were extracted from the register. Specifically, the patient's sex, ethnicity, age at diagnosis, full address at diagnosis, and date of diagnosis were obtained (Table 5.1). The full address at diagnosis was requested for geocoding purposes, and the use of these data was approved by the Ethics Committee of the local health authority (Upper South A).

Geocoding

The addresses were first run through a batch geocoder and any remaining unmatched addresses were checked for spelling/number/suburb mistakes. The internet mapping services, Google Maps (Google, 2007), and ZoomIn (ProjectX Technology Ltd., 2006) were used to check the addresses for accuracy. 87.5 percent of cases were coded as exact matches by the programme, and each of these matches was checked manually, with no mistakes noted. For some of the remaining unmatched addresses, the correct street could be located, but the street number was not found in the address database. This occurrence has been noted in other New Zealand studies (e.g. Skelly et al., 2002) and is likely to be a result of property subdivision or new housing developments. In this instance, cases were coded to the nearest recognised street number. After adjusting minor spelling mistakes and changing unrecognised street numbers to a nearby number, the figure for exact matches increased to 95.5 percent. Again, all new matches were checked for accuracy. Of the remaining 4.5 percent; 5 cases were mapped to the correct street, but possibly not the correct CAU since these streets were long and therefore dissected more than one CAU; 7 were mapped to a rural township, and 3 cases were mapped to PO Box addresses. It was not known how close each patient lived to the PO Box where their post was delivered.

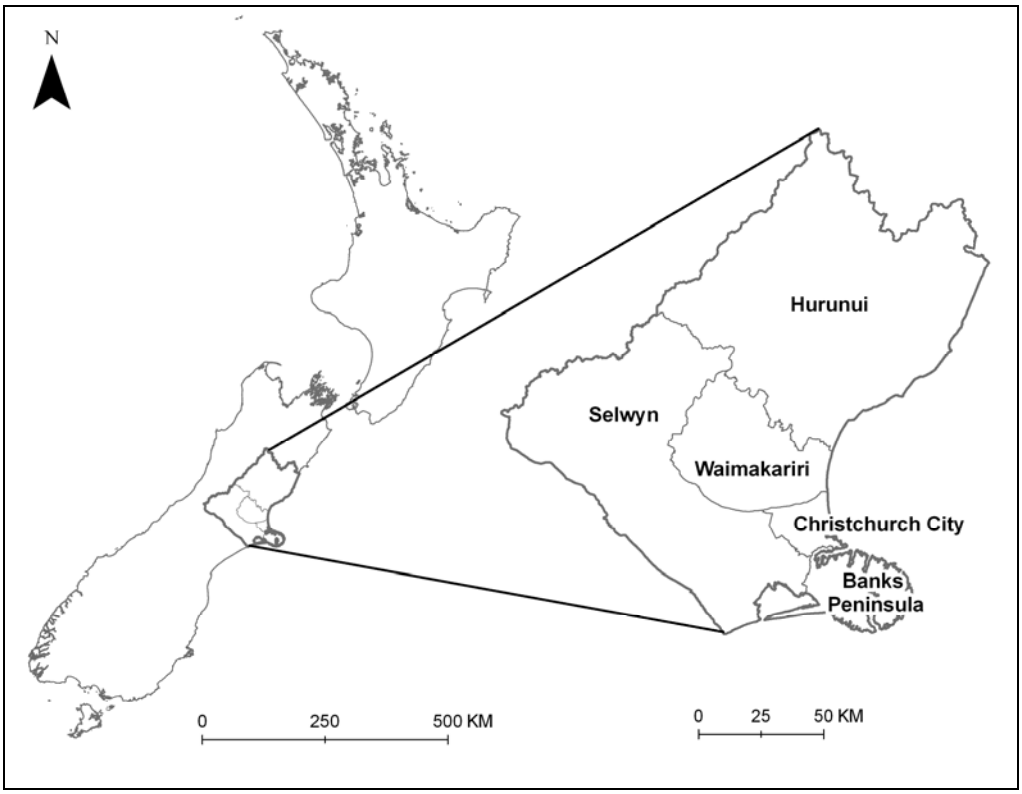

Figure 5.1: Coverage of the Canterbury Diabetes Register

5.3 Methods

5.3.1 Geographical epidemiology of ALL and type 1 diabetes

The first aim of this research was to determine the geographical epidemiology of childhood ALL and type 1 diabetes in New Zealand. A number of different methods were employed. First, the descriptive epidemiology of each disease was determined at the national-level using incidence proportions. Second, spatial and temporal patterns of both diseases were examined at a more detailed spatial scale, using standardised incidence ratios, Poisson probabilities and cluster analyses.

5.3.1.1 Incidence proportions

Differences in the risk of developing ALL or type 1 diabetes have been shown to depend upon a child's age, sex and ethnicity (Chapter 3 and 4). In order to assess how the disease burden at the national-level varied by these individual risk factors, incidence proportions were calculated. Incidence proportions were calculated for various age groups at diagnosis (0-4, 5-9 and 10-14 years), ethnic groups (European, Māori, Pacific and Asian) and each sex separately. Incidence

proportions were calculated as the number of new cases which occurred in each subgroup during the study period (1980-2004), divided by the total population at risk. The resulting figure was multiplied by 100,000 to express the number of new cases occurring per 100,000 people at risk (Moon et al., 2000). Population data from the 1991 census were employed to enumerate the total population at risk, since this year was the approximate mid-point of the study. The population measured at the mid-point of a study period has been considered an adequate estimation of the population at risk for relatively rare diseases (Breslow and Day, 1980). Incidence proportions were also calculated by the year of diagnosis for preliminary assessment of how the incidences of these diseases have changed over time in New Zealand. Such incidence proportions were also disaggregated by age group at diagnosis and sex, with three year moving average incidence proportions calculated where the numbers were small. This method removes short term fluctuations in incidence in order to reveal overall temporal trends.

5.3.1.2 Standardised incidence ratios (SIRs)

Introduction

While determining national-level incidence proportions is useful for comparison with other countries, it is also important to assess the variation in incidence at smaller spatial scales. Small area variation in childhood ALL (e.g. Alexander et al., 1996, Feltbower et al., 2005) and type 1 diabetes (e.g. McKinney et al., 1996, Staines et al., 1997) incidence have been noted in the UK, but not previously investigated in New Zealand. However, mapping the raw counts or crude incidence proportions of diseases can be misleading as such measures do not take into account differences in the structure of the population in each area. One way to take into account the age and sex structure of each population is to compare age- and sex-specific incidence rates. However, this method becomes cumbersome when comparing many age and sex groups for different areas (Moon et al., 2000). Moreover, the populations may not be sufficiently large or the disease common enough for such rates to be reliable (MacMahon and Trichopoulos, 1996). An alternative method involves the calculation of a single rate that adjusts for the age and sex differences between populations. This process is termed standardisation, of which there are two main methods: direct and indirect standardisation (Webb et al., 2005).

Direct standardisation (for age) involves calculating the overall incidence rate that would be expected to be found in a standard population, if it had the same age-specific rates as the study population (Webb et al., 2005). Direct standardisation is preferable to indirect standardisation as

it is able to adjust precisely for the affects of age (MacMahon and Trichopoulos, 1996). However, unstable estimates occur where the number of cases in each age group in the study population are small (Moon et al., 2000). Cases of childhood ALL and type 1 diabetes by age group at diagnosis for the whole study period (1980-2004) were small at the national and regional level (Table 5.1). Thus, disaggregating age-specific cases for smaller geographies such as CAUs would have produced highly unstable rates. It is for this reason that the method of indirect standardisation was employed.

In indirect standardisation, the number of observed cases in a study area is compared to the number of cases that would have been expected had the incidence rates in the study population been the same as those for the standard population (Webb et al., 2005). First, an age-specific incidence rate (ASIR) is calculated for each age group in a chosen standard population (usually the national rates for the country or area in which the study was undertaken). These ASIRs are then multiplied by the populations in the corresponding age groups in the study area, and summed to give the expected number of cases in the study area if it had the same ASIRs as the standard population. The actual observed number of cases in the study area is then divided by the expected number of cases, and multiplied by 100 to produce a standardised incidence ratio (SIR) (MacMahon and Trichopoulos, 1996, Pringle, 2003). If the SIR is equal to 100, the study population had the same risk of disease as the standard population, after correction for differences in age. The observed number of cases is equal to that expected. If the SIR is greater than 100, then the study area's incidence is greater than that of the standard population. If it is less than 100, then the study area's incidence is lower than that of the standard population (Moon et al., 2000).

SIRs by CAU and TA

In this study, age SIRs were calculated at the TA and CAU-level, using data for the whole study period (1980-2004). Due to the slightly larger number of type 1 diabetes cases at the CAU-level, SIRs for this disease were also calculated for the periods 1980-1991 and 1992-2004 to examine changes over time. The standard population was defined as the whole of New Zealand for the purposes of the ALL SIR calculations, and the whole of Canterbury for the type 1 diabetes calculations. The ASIRs were computed using population data at the approximate mid-point of the study period being analysed (Tables 5.2 and 5.3). The resulting SIRs were mapped in geographical information systems (GIS) software to identify areas which had higher or lower than expected rates of each disease.

In order to assess the reliability of the SIRs, the significance of each was tested using chi-square values (χ^2), and confidence intervals (CIs) were computed to define the range within which the true SIR value was likely to lie. The χ^2 value was calculated as:

$$\chi^2 = \frac{(O-E)^2}{E}$$

Where O is the observed number of cases and E is the expected number of cases in each CAU/TA. The larger the χ^2 value, the greater the difference between the observed and expected values. More specifically, if the χ^2 value exceeds 3.8, the SIR is considered to be statistically significant (Moon et al., 2000). Where the observed number of cases was greater than 50 (TA-level analyses only), the 95 percent confidence intervals were calculated as:

= SIR +/- 1.96*SE

Where SE is the standard error calculated as the SIR divided by the square root of the expected number of cases (Moon et al., 2000). However, where the number of observed cases was less than 50, the Poisson distribution was used to calculate exact confidence intervals:

$$\text{Lower confidence interval (LCI)} = \frac{SIR}{O} * \mu_0$$

$$\text{Upper confidence interval (UCI)} = \frac{SIR}{O} * \mu_1$$

Where μ_0 and μ_1 are the exact confidence intervals for the expectation (μ) of a Poisson distribution according to the number of observed cases (O) (Esteve et al., 1994, p.61).

Table 5.2: Age-specific incidence rates of ALL per 100,000 in New Zealand 1980-2004

Age group	**Total ALL cases**	**Population 1991**	**Age-specific incidence rate (ASIR)**
0-4 years	436	277,185	157.30
5-9 years	207	251,115	82.43
10-14 years	138	255,300	54.05

Table 5.3: Age-specific incidence rates of type 1 diabetes per 100,000 in Canterbury 1980-2004

Age group	**Total type 1 diabetes cases**	**Population 1991**	**Age-specific incidence rate (ASIR)**
0-4 years	78	24,975	312.31
5-9 years	101	22,905	440.95
10-14 years	158	23,898	661.14

SIRs by area-type

Previous studies have also shown both ALL and type 1 diabetes incidence to vary by different types of area. For example, both diseases have been associated with area-level deprivation, with the majority of studies revealing an inverse association (e.g. Draper et al., 1991, Haynes et al., 2006, Parslow et al., 2001, Poole et al., 2006). However, there are some exceptions to this general rule (e.g. du Prel et al., 2007) and the potential relationships between area-level deprivation and either disease have not previously been explored in a New Zealand context. Consequently, the numbers of observed and expected cases in each CAU (calculated as above) were aggregated into deprivation categories.

In New Zealand, area-level deprivation has been measured using the New Zealand Deprivation Index 2001 (NZDep01). The NZDep01 combines nine variables from the 2001 census which reflect eight dimensions of social deprivation (see Table 5.4) (Salmond and Crampton, 2002a). Principal components analysis was used to create a deprivation score from these variables for each meshblock area in New Zealand (Crampton et al., 1997). From these data, weighted average NZDep01 scores were also calculated for CAUs (Salmond and Crampton, 2002b). The NZDep01 ten point scale (NZDep01 deciles) and NZDep01 five point scale (NZDep01 quintiles) were derived from this interval variable. For example, NZDep01 deciles divide New Zealand into tenths of the distribution of the first principle components scores. As a result, a CAU with a deprivation decile of 10 indicates that it is in the most deprived 10 percent of CAUs in New Zealand (Figure 5.2) (Salmond and Crampton, 2002a). Thus SIRs were calculated for deprivation deciles and quintiles at the CAU-level.

Table 5.4: NZDep01 variables (Source: Salmond and Crampton, 2002a)

Dimension of deprivation	Variable description (in order of decreasing weight)
Income	People aged 18-59 receiving a means tested benefit
Employment	People aged 18-59 unemployed
Income	People living in equivalised* households with income below an income threshold
Communication	People with no access to a telephone
Transport	People with no access to a car
Support	People aged <60 living in a single parent family
Qualifications	People aged 18-59 without any qualifications
Owned home	People not living in own home
Living space	People living in equivalised* households below a bedroom occupancy threshold

* Equivalisation: methods used to control for household composition

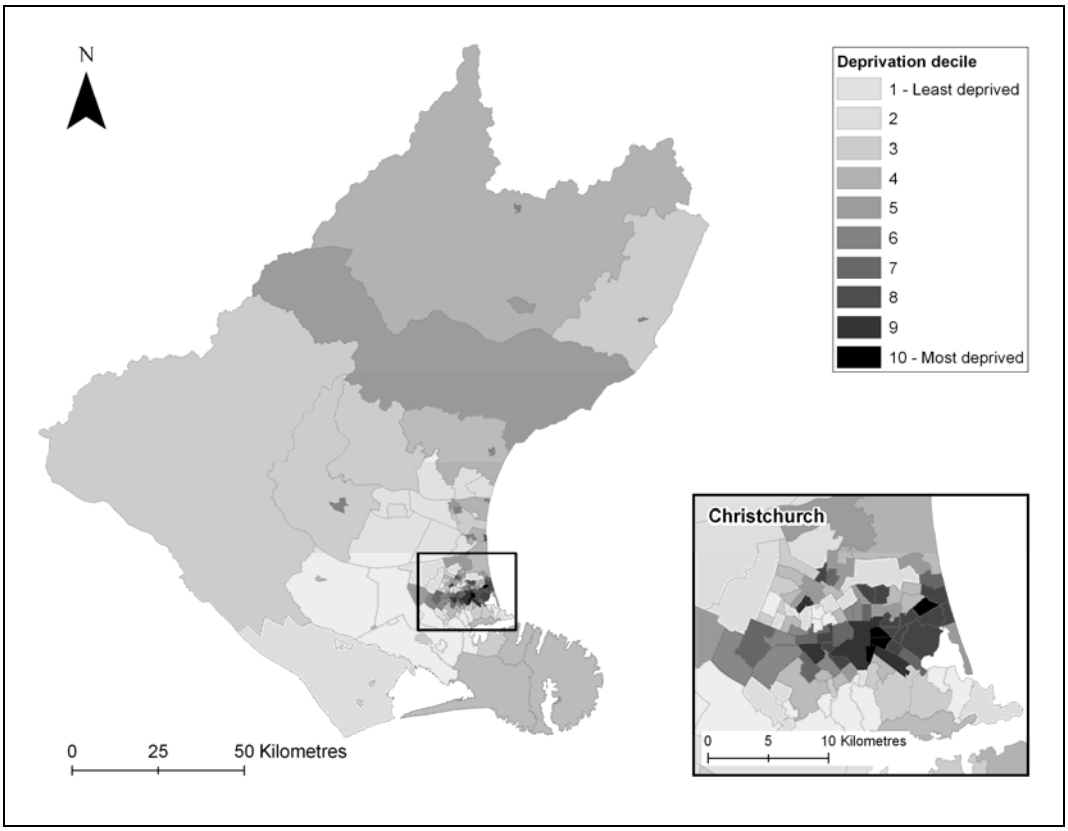

Figure 5.2: CAUs in the Canterbury diabetes study region by deprivation decile 2001

As well as variations by area-level deprivation, overseas studies have shown the incidence of both diseases to differ by the urban/rural nature of areas (e.g. Adelman et al., 2005, Li et al., 1998, Rytkonen et al., 2003, Waugh, 1986). However, the associations found vary depending on

the geographical setting and the type of classification used, and to date no attempt has been made to examine this issue in New Zealand. Consequently, the numbers of observed and expected cases by each CAU were also summed according to Statistics New Zealand's urban/rural classification of CAUs. This profile assigns all CAUs across the country into one of seven categories: main urban areas, satellite urban communities, independent urban communities, rural areas with high, moderate and low urban influence, and highly rural/remote locations (Figures 5.3 and 5.4). This classification has the advantage over other urban/rural categorisation methods (such as population density) of being based on workplace address relative to home address, and therefore allows economic and social ties between urban and rural areas to be incorporated into the classification (Statistics New Zealand, 2006a).

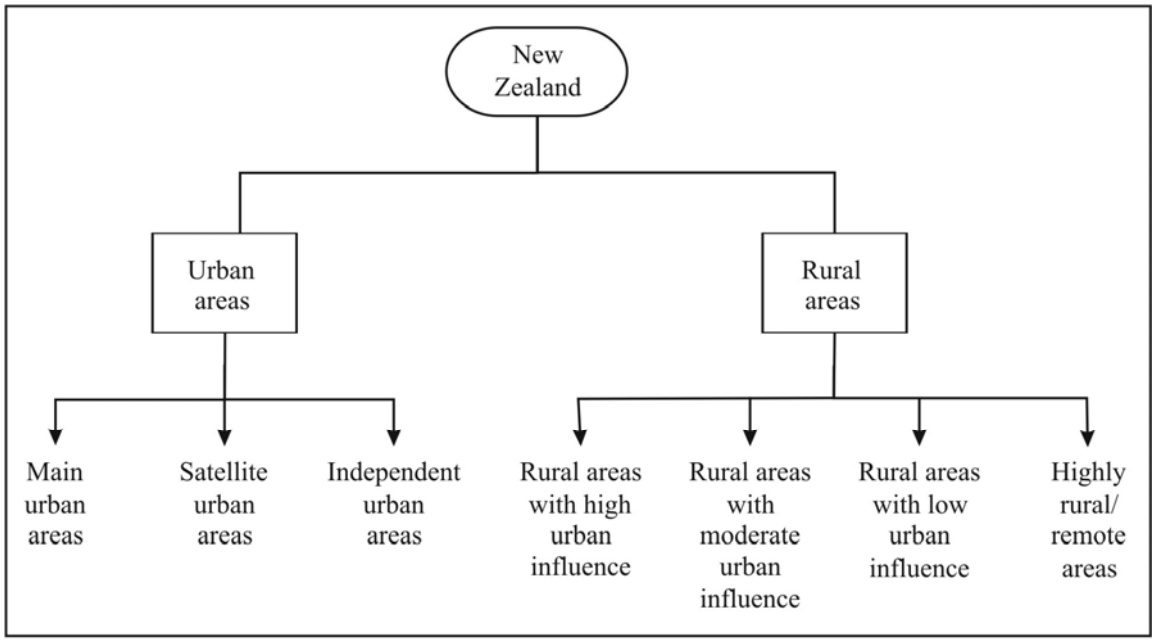

Figure 5.3: Urban/rural classification of CAUs in New Zealand
(Adapted from: Statistics New Zealand, 2006a)

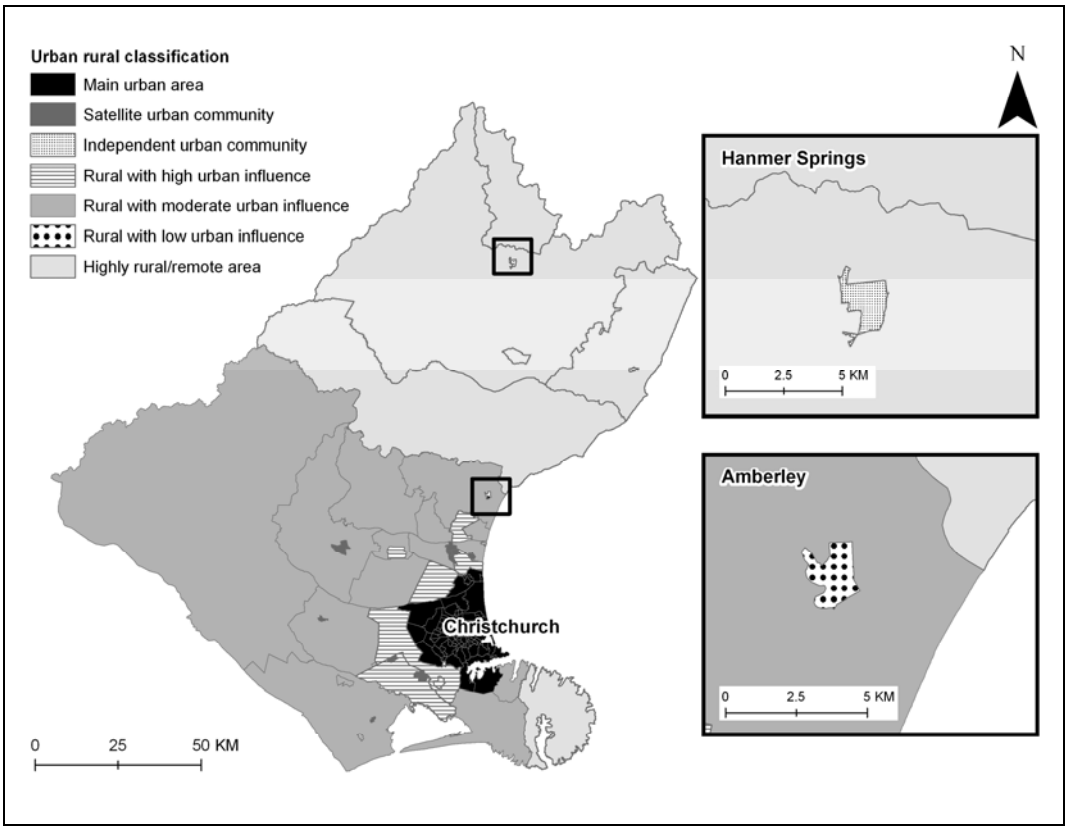

Figure 5.4: CAUs in the Canterbury diabetes study area by urban/rural classification 2001, with insets of the only independent urban community (Hamner Springs) and the only rural area with a low urban influence (Amberley) in this region

A major drawback of maps of SIRs and even SIRs calculated for different area-level categories is that the smaller the underlying population at risk, the more the rate estimates are influenced by random variability. An alternative option is to produce maps based on p-values from tests of whether the rate in each area differs significantly from the overall rate (Olsen et al., 1996)

5.3.1.3 Poisson probabilities

The confidence intervals of SIRs can help to constrain their values, but when small numbers are involved it is preferable to re-examine the observed and expected frequencies in light of an appropriate statistical model (Giles, 1983). Given that both diseases are rare in the general population, it is reasonable to test their observed occurrence against a random expectation based on the Poisson distribution. The Poisson probability distribution describes the likelihood of rare, random events in a continuum, given a mean expectation and a variance equal to the mean (Long, 1997). Figures 5.5 and 5.6 show, that for both diseases, the Poisson distribution represents a reasonable fit to these data. A comparison of observed values with those predicted by the Poisson generating function for each CAU gives a measure of the likelihood of the

observed values occurring by chance. The probability (P) of an event (x) occurring can be calculated as:

$$P(x) = e^{-a} \frac{a^x}{x!}$$

Where: a = the mean number of cases for all areas

x = the observed number of cases for a particular CAU

e = the constant, 2.7183

An improvement can be made to this equation by replacing the mean incidence of disease for all areas (a) with the expected number of cases for individual CAUs (E):

$$E = \sum mx\ \lambda x$$

Where λx is the incidence rate of group x in the standard population (the ASIR) and mx is the number of person years accumulated by group x in the population under study. Low Poisson probabilities indicate a larger observed than expected value (Pearce and Boyle, 2005). For example, a Poisson probability of 0.01 indicates that there is a significant difference between the observed and expected values at the 99 percent confidence level. The Poisson probabilities for each disease were mapped in GIS software, with large circles representing the CAUs with the greatest difference between the expected and observed values.

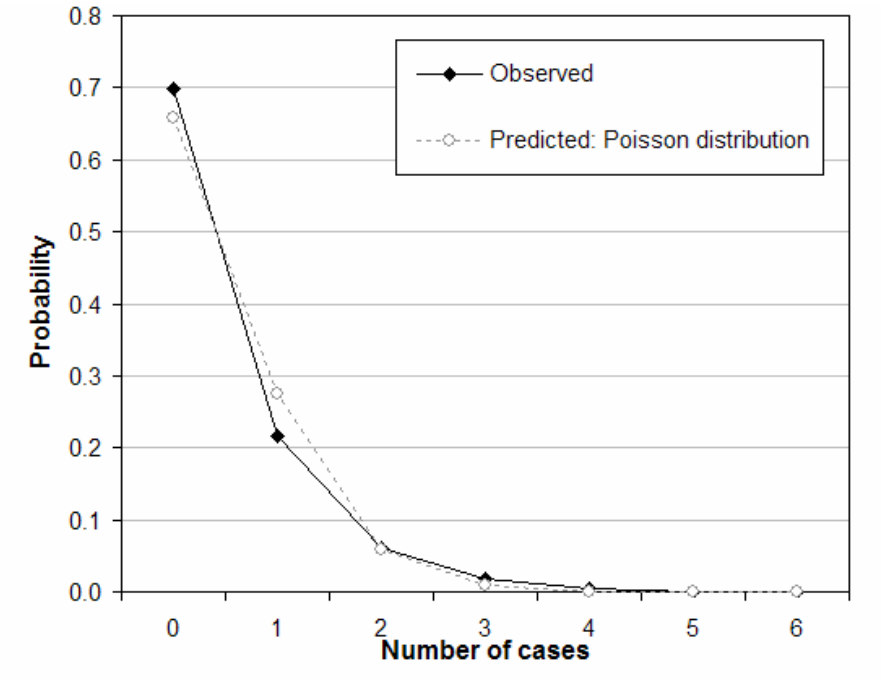

Figure 5.5: Distribution of observed and predicted ALL cases at the CAU-level in New Zealand 1980-2004

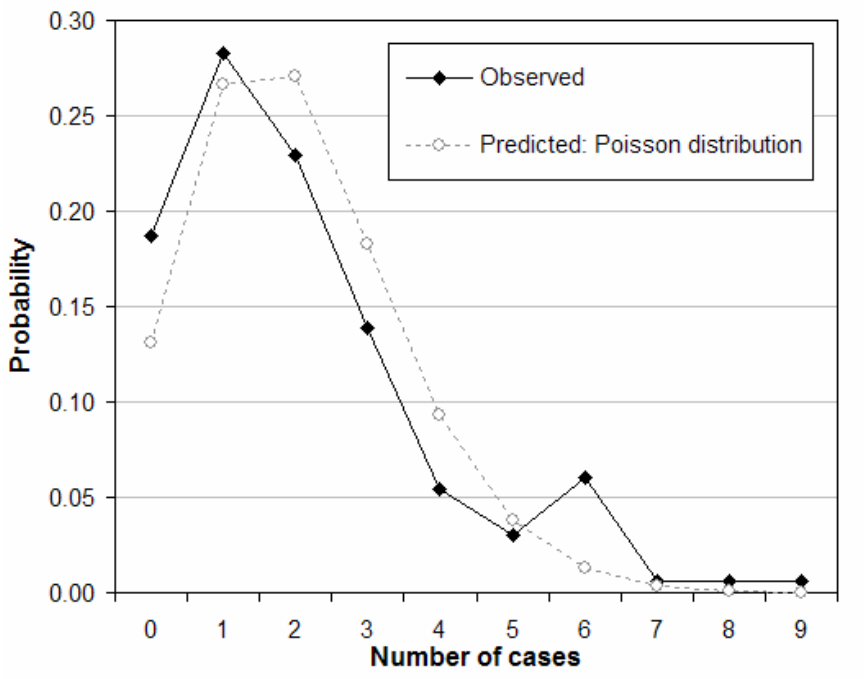

Figure 5.6: Distribution of observed and predicted type 1 diabetes cases at the CAU-level in Canterbury 1980-2004

5.3.1.4 Cluster analysis

Background

The question of whether diseases are clustered in time and space has generated considerable interest within public health (Lawson, 2001, Tango and Takahashi, 2005). A disease cluster is where more cases are identified within a certain group of people, geographic area, and time period than are expected based on the size and age of the population (Thun and Sinks, 2004). More specifically, a spatial cluster can be defined as a geographically bounded group of occurrences of sufficient size and concentration to be unlikely to have occurred by chance (Knox, 1989). Cluster analyses in epidemiology can be used to detect the aggregation of disease cases, to test whether there is any true clustering and most importantly, to generate hypotheses regarding disease aetiologies (Hjalmars et al., 1996, Olsen et al., 1996).

Significant clusters of both childhood leukaemia (Chapter 3) and type 1 diabetes (Chapter 4) have been reported in several settings. Moreover, two early studies in New Zealand noted significant clustering of leukaemia in children aged less than six years. However, to date, no studies in New Zealand have tested for clustering of childhood type 1 diabetes. Moreover, since the early leukaemia cluster analyses were conducted in New Zealand, many new statistical methods have been developed to test for clusters. However, many methods still suffer from multiple testing problems as a result of the requirement for one or more unknown parameters to be defined prior to analysis (Tango and Takahashi, 2005). In addition, most statistical methods for spatial cluster analysis are either able to descriptively find the location of clusters without any way of determining their significance, or are able to test whether there are any statistically significant clusters without being able to locate them spatially (Kulldorff, 1997). For example, the most recent New Zealand study employed the Cuzick and Edward's method of cluster detection which tests for clustering throughout the study region without the ability to pinpoint the location of specific clusters. Furthermore, this method is not very powerful at detecting small clusters existing for short periods of time (Dockerty et al., 1999c). The spatial scan statistic in Satscan software was employed in this research as it has the ability to both spatially locate clusters and provide tests of statistical significance. Moreover it overcomes the problem of multiple hypothesis testing as the software allows the scanning circle size to vary continuously (Kulldorff, 2006).

In Satscan, the space-time scan statistic uses a cylindrical window with a circular (or elliptic) base which corresponds to geographical area and height which corresponds to time. This window is then moved over the study area, in space and time, so that for each possible geographical location and size, it also scans through each possible time period and size. An important advantage of this method is that it is able to detect clusters without prior knowledge of their location, geographical size, when they might occur or how quickly they might emerge. The p-value is obtained through Monte Carlo hypothesis testing which involves comparing the rank of the maximum likelihood from the real data set with the maximum likelihoods from the random data set (Kulldorff, 2006). For 9,999 replications the test can be considered significant at the 5 percent level if the value of the test statistic for the real data set is among the 500 highest values of the test statistic from the replications.

Analysis

Since both diseases are relatively rare outcomes, the Poisson model, in which the number of cases in each location is assumed to be Poisson distributed (Kulldorff, 1997), was employed. The analyses were run at two different geographical scales, the CAU and meshblock-level to examine whether scale had an important effect on the results achieved. A meshblock (MB) is the smallest data collection unit used in New Zealand; rural MBs have an average population of 60 people, compared to 110 people in urban areas (Statistics New Zealand, 2007b). In order to reduce computation time, separate runs for the North and South Island were carried out for the ALL MB analyses. Each case was assigned to a CAU/MB based on their address at diagnosis, and the seven digit New Zealand Map Grid coordinates for the centroid of each CAU/MB in New Zealand were used in the analyses.

As well as being run at different geographical scales, the analyses were also divided by age group, since rates of both diseases have been shown to vary considerably in under 15 year olds (Little, 1999, Slama, 2003). It is conceivable that different causal mechanisms or pathways are at work for children of different ages which may be discernible in cluster analyses. As a result, the analysis was run initially for all children aged 0-14 years, and then for those aged 0-4 years, 5-9 years and 10-14 years (Table 5.5).

Sex differences have also been noted for both diseases (Karvonen et al., 1997a, Liang and Pui, 2005). Sex differences were only tested at the CAU-level due to the prohibitively long computation times associated with MB analyses. For ALL, sex differences were only tested at

the CAU-level for all age children, as these analyses included CAUs for the whole of New Zealand. Since the type 1 diabetes data were only available for the Canterbury region, running times for these analyses were much shorter, and as a result, cases were also disaggregated by age group and sex for the CAU-level analyses (Table 5.5).

Since no covariates were included, the expected number of cases was proportional to the age group- and sex-specific population count for each area. For the CAU-level analyses, population data by age group and sex were obtained from the Ministry of Health for the years 1981 and 2001. These years were chosen as they were census years at the beginning and end of the study period. Where two years of population data are provided, a linear interpolation of the population numbers between the start and end points was calculated. For the MB analyses, population count data were only available after 1991 from Statistics New Zealand. For the type 1 diabetes analyses, data from the 1991 and 2001 censuses were included. For the ALL analyses, the population count data were taken from the 1991 census only, as adding further years of population data was computationally infeasible with current computer processing speeds. As a result, population counts from the 1991 census were utilised as this year represented the approximate mid-point of the study period. Population counts of zero were recoded to one to enable the running of the analysis.

Table 5.5: Cluster analyses conducted 1980-2004

Analyses	ALL New Zealand	Type 1 diabetes Canterbury
MB		
All cases	✓	✓
0-4 years	✓	✓
5-9 years	✓	✓
10-14 years	✓	✓
CAU		
All cases	✓	✓
0-4 years	✓	✓
5-9 years	✓	✓
10-14 years	✓	✓
Male	✓	✓
Female	✓	✓
0-4 years male	-	✓
0-4 years female	-	✓
5-9 years male	-	✓
5-9 years female	-	✓
10-14 years male	-	✓
10-14 years female	-	✓

The programme scans for clusters of geographic size between zero and a pre-defined upper limit. A 50 percent scanning window is recommended as this size allows both small and large clusters to be detected without any pre-selection bias in terms of cluster size (Kulldorff, 2006). Thus, for the CAU-level analyses the maximum spatial cluster size was set at 50 percent of the population at risk, and the maximum temporal cluster size was set at the recommended level of 50 percent of the study period (12.5 years), for both diseases (Table 5.6). However, for the MB-level analyses it was not possible to use such a large upper limit for the scanning window due to the increased processing power required for scanning over 7,400 MBs in Canterbury and over 38,000 MBs in the whole of New Zealand. MB-level analyses were therefore conducted using smaller windows (5 percent for ALL or 25 percent for Type 1 diabetes) and were thus useful for detecting relatively small clusters.

Table 5.6: Cluster analyses: scanning window sizes

Analyses	**Scanning window size**	
	ALL ***New Zealand***	***Type 1 diabetes*** ***Canterbury***
MB		
All cases	5%	25%
0-4 years	5%	25%
5-9 years	5%	25%
10-14 years	5%	25%
CAU		
All cases	50%	50%
0-4 years	50%	50%
5-9 years	50%	50%
10-14 years	50%	50%
Male	50%	50%
Female	50%	50%
0-4 years male	-	50%
0-4 years female	-	50%
5-9 years male	-	50%
5-9 years female	-	50%
10-14 years male	-	50%
10-14 years female	-	50%

5.3.2 Examining the relationship between population mixing and ALL/type 1 diabetes

The methods described thus far were employed to determine both the descriptive and geographical epidemiology of childhood ALL and type 1 diabetes in New Zealand/Canterbury. Therefore, these methods addressed the first aim of the thesis. The second aim of this work was to explore appropriate ways of measuring population mixing in New Zealand, and to examine whether population mixing is associated with the incidence of childhood ALL and/or type 1

diabetes in this context. Consequently, the following sections detail how population mixing was quantified, and subsequently analysed in relation to both diseases.

5.3.2.1 Measuring population mixing

In Chapter 2 population mixing was defined as the movement and interaction of people over time and space. Movements of people were therefore measured (interaction could only be assumed) and used as a proxy for area-level infectious exposure. Population mixing was measured at the area-level since the spread of infections is governed by herd immunity, which cannot be usefully measured at the individual-level (John and Samuel, 2000, Kinlen et al., 1990).

Contrary to a number of previous studies, this research did not focus upon population mixing occurring in remote rural areas. Population mixing levels were ascertained for CAUs across the whole of New Zealand as excesses of childhood leukaemia and type 1 diabetes cases have been observed in several metropolitan/urban settings. Moreover, significantly raised childhood leukaemia has been noted in urban wards with high population mixing levels in England and Wales (Dickinson et al., 2002). Additionally, social isolation (rather than just geographical isolation) can occur anywhere, and also has consequences for the spread of infections.

Previous studies have suggested that a range of different population movements can affect the transmission of infections, and as a result, have implications for the incidence of childhood ALL and/or type 1 diabetes. For example, the importance of population growth, the volume of in-migrants, the distance travelled by in-migrants, and the diversity of their origins have all been found to be significantly related to either or both childhood ALL or type 1 diabetes (Chapter 2). As a result, a number of different measures were created in order to account for the volume, frequency, distance and diversity of population movements into and within New Zealand.

Importantly in terms of the population mixing objectives of this thesis, static population mixing measures were created as an estimate of early life infectious exposure in order to test the hygiene hypothesis relating to childhood type 1 diabetes; and changes in population mixing measures over time were created in order to test whether increases were associated with triggering either childhood ALL or type 1 diabetes. Initially, a number of static (or snapshot) population mixing measures were created for every CAU in New Zealand for each census year (1981, 1986, 1991, 1996 and 2001) (Table 5.7).

Static population mixing measures

First, the number of total migrants was calculated as a percentage of the whole population in each CAU. Migrants were defined as people whose usual residence five years ago was either elsewhere in New Zealand (internal movers) or overseas (international movers). The total migrants' variable included movers of all ages. This measure gives an indication of the number of new people moving into an area, and also how much of the total population they account for.

Migrants were then differentiated by age (child migrants: 5-14 years), and distance travelled (overseas migrants) to test theories that childhood population mixing may be more important in the pathogenesis of both diseases (Feltbower et al., 2005, Parslow et al., 2001) and that migrants from further away are likely to bring infections that pose a greater immunological challenge to local children (Kinlen et al., 1990). Since the migrant variables (total, child and overseas) enumerated movements of people in the previous five years, a one year mobility percentage was also computed to capture shorter term residential relocations. This measure detailed the percentage of people who had been in their current residence for less than one year.

As noted in Chapter 2, a large number of new residents in an area may not be enough to trigger excesses of childhood ALL or type 1 diabetes; the diversity of their origins is also likely to be important. As a result, the Shannon Index of Diversity was employed to gauge whether migrants came from a diverse or narrow set of origins (Stiller and Boyle, 1996). This measure was originally used in ecology to quantify both the richness of species in an area (e.g. the total number of species) and the evenness of these species (e.g. the distribution of individual plants in each species) (Magurran, 1988). Applying the measure to migration research, the index measures the extent to which the total incomers to an area are distributed among their origin areas. Generally, the greater the number of origin areas, the greater the number and range of infections that can be introduced, and the more likely it is that there will be appreciable differences in herd immunity within the area (Kinlen et al., 1990). This aspect of the equation deals with the richness of in-migration. However, evenness is important too. For example, if a TA receives 100 migrants from 5 origin areas, but 96 of these migrants all come from the same origin area; a child would be less likely to come into contact with one of the other 4 migrants. The Shannon Index of Diversity takes this aspect into account.

In this study, migrant diversity (H) was calculated for every CAU in New Zealand as:

$$H = - \Sigma (n_i/N) \ln (n_i/N)$$

Where n_i is the number of migrants from ith territorial authority (TA) area (excluding the TA which the CAU falls within) or overseas country, and N is the total number of migrants in the CAU (Table 5.8). Other studies have included a correction to this formula to account for under-ascertainment or error in census data (e.g. Law et al., 2003). However, it has been noted that the error is rarely significant (Magurran, 1988, Parslow et al., 2001) and thus the correction was not included in this research. The value of the Shannon Index of Diversity is usually found to fall between 1.5 and 3.5 and rarely surpasses 4.5 (Magurran, 1988). Higher values indicate a higher diversity of origins of incoming migrants.

Recent work by Cox (2007) in Scotland, argued that various combinations of in-migration and migrant diversity will have different implications for the spread of infection. He consequently created a new categorical variable using both these measures. Following his method, the mean value of the percentage of total migrants and the mean value of migrant diversity were firstly calculated for each census year. Then, for every census year, each CAU was assigned to a category of either high or low migration depending on whether the value for the percentage of total migrants was above or below the New Zealand mean. Each CAU was also classified as having either high or low migrant diversity, depending on whether its migrant diversity value was above or below the national average. The two categories were then merged to form four new categories:

1. Low in-migration and low diversity
2. Low in-migration and high diversity
3. High in-migration and low diversity
4. High in-migration and high diversity

Finally, the percentage of overseas visitors was also derived to establish whether New Zealand's increasing number of tourists (Statistics New Zealand, 2006b) are important in the aetiology of childhood ALL or type 1 diabetes. Movements of tourists have yet to be examined in population mixing and health studies (Miller et al., 2007). In New Zealand, this omission is crucial as visitors from overseas are highly mobile within the country, and visit relatively isolated areas

(Ministry of Tourism, 2007). In places where tourists make up a large percentage of the total population, contact with permanent residents and their children are more likely.

The numbers of movers (migrants or visitors) were calculated as a percentage of the total population in each area in order to establish how prevalent they were among the general population, and to aid comparison between areas of differing population size. It is argued that where movers make up a high percentage of the population in an area, contact with a susceptible child is more likely, and contact is essential for the potential spread of infections (Anderson, 1982).

Table 5.7: Static population mixing measures

Population mixing measure	Description	Justification	Used previously?
Percentage of total migrants	% of people whose usual residence 5 years ago was either elsewhere in New Zealand (internal movers) or overseas (international movers)	Describes the volume of 'new' people in an area. Has implications for the number of infections brought in	Yes (e.g. Parslow et al., 2001, Rudant et al., 2006, Stiller and Boyle, 1996)
Percentage of child migrants	% of children (5-14 years) whose usual residence 5 years ago was either elsewhere in New Zealand (internal movers) or overseas (international movers)	As above, but for children only. It has been hypothesised that child mixing may be more important for both diseases	Yes (e.g. Law et al., 2003, Parslow et al., 2001, Stiller and Boyle, 1996)
Percentage of overseas migrants	% of people whose usual residence 5 years ago was overseas (international movers)	As above, but for overseas migrants only. People from overseas may bring new infections into an area	Yes (e.g. Cox, 2007, Dickinson et al., 2002)
One year mobility percentage	% of people who have been in their current residence for less than one year	To give an indication of more recent moves (within the last year) since the migrant variables only detail moves since the last census (5 years ago)	No
Migrant diversity score	Diversity of origin of all migrants; internal (excluding those moving within the same TA) and overseas migrants	The importance of newcomers from a range of different origins has been associated with both diseases & thought to have implications for the range of infections introduced	Yes (e.g. Feltbower et al., 2005, Law et al., 2003, Parslow et al., 2002, Stiller and Boyle, 1996)
Percentage of overseas visitors	% of tourists who were present in New Zealand on census night	To test the role of overseas tourist frequencies in the aetiology of both diseases	No
Population mixing category	Categorisation of areas based on whether their % of total migrants was above (high) or below (low) average, and whether their migrant diversity score was above or below average for that census year:	As argued by Cox (2007), different combinations of in-migration and migrant diversity together, could have varying implications for infection spread	Yes (Cox, 2007)
1. Low in-migration & low migrant diversity	areas with below average in-migration & migrant diversity		
2. Low in-migration & high migrant diversity	areas with below average in-migration & above average migrant diversity		
3. High in-migration & low migrant diversity	areas with above average in-migration & below average migrant diversity		
4. High in-migration & high migrant diversity	areas with above average in-migration & migrant diversity		

Table 5.8: Example of migrant diversity score calculations

	Number of migrants from each territorial authority / overseas country of origin						***Shannon Index of Diversity (H)***
CAU	**Far North District**	**Whangarei District**	**Kaipara District**	**Rodney District**	**North Shore City**	**etc….**	
514000	9	18	6	36	171	…	3.0631
514101	9	12	3	15	96	…	2.8898
514102	9	24	3	36	168	…	2.9574
514103	21	33	6	51	156	…	3.2031
514200	3	9	0	6	48	…	2.8283
514301	3	21	3	12	45	…	3.1329
514302	6	15	0	6	30	…	3.1527
514401	12	9	6	24	42	…	2.5664
etc…	…	…	…	…	…	…	…

Change over time

Since this research also aimed to test the potential triggering role of an increase in population mixing over short periods of time, the change in each of the static population mixing measures was also calculated. For every CAU in New Zealand, the percentage change in each measure was calculated for the years: 1981-2001, 1981-1991, 1991-2001, 1981-1986, 1986-1991, 1991-1996 and 1996-2001.

Population mixing (PM) change was calculated as (Miller, 2005):

$$\frac{(\text{PM year 2} - \text{PM year 1})}{\text{PM year 1}} * 100$$

In addition, the percentage change in the total population of each CAU relative to the first year was also calculated. While this measure was critiqued in Chapter 2, it was included in these analyses to aid comparisons with previous studies (e.g. Dockerty et al., 1996a, Koushik et al., 2001, Langford, 1991, Wartenberg et al., 2004) and with the more direct measures of in-migration also used in this research.

In order to see whether the population mixing categories of each area changed over time (between year 1 and year 2), all of the possible combinations of change were first listed in a table (columns 1-3, Table 5.9). From the derived descriptions of change, a new set of categories were formed (columns 4 and 5, Table 5.9). The first category (1) included areas where either/both in-

migration and migrant diversity decreased to below the national average by year 2. Decreases in population mixing were assumed to result in a lower number and/or range of potential infections circulating in an area and therefore not likely to trigger higher incidence of ALL or type 1 diabetes. As a result, these areas were not further defined for the purposes of this research, and were used as the base category in the analyses. The second category (2) included areas where the population mixing categories were the same in both years. Thus in-migration and migrant diversity both remained above or below the national average by year 2. Since these areas did not witness any above average increases in either in-migration or migrant diversity, these were not expected to be associated with raised incidence of the diseases and were thus not further defined. Category 3 and 4 areas were those which experienced an increase in either in-migration and/or migrant diversity. In detail, category 3 areas included CAUs where migrant diversity increased from below to above average by the second year, but which witnessed no substantial changes in in-migration. Category 4 areas included CAUs where in-migration increased to above the national average, and migrant diversity either remained the same increased or decreased. Category 4 areas were expected to be associated with increased incidence of ALL and type 1 diabetes. The volume of in-migration was considered to be more important for disease causation than migrant diversity since if areas have a high migrant diversity score but these migrants only make up a small proportion of the total population, they are less likely to come into contact with susceptible children.

Table 5.9: Population mixing (PM) change category workings

All possible PM category combinations			New Categories	
Year 1 PM category	**Year 2 PM category**	**Description of change from year 1 to year 2**	**Description**	**Number**
LMLD	HMHD	LMLD to HMHD	Increase in migration category	4
LMLD	HMLD	LMLD to HMLD	Increase in migration category	4
LMLD	LMHD	LMLD to LMHD	Increase in diversity category	3
LMLD	LMLD	No Change	No change in either category	2
LMLD	NO DATA	Decrease	Decrease in either/both categories	1
LMHD	HMHD	LMHD to HMHD	Increase in migration category	4
LMHD	HMLD	LMHD to HMLD	Increase in migration category	4
LMHD	LMHD	No Change	No change in either category	2
LMHD	LMLD	Decrease	Decrease in either/both categories	1
LMHD	NO DATA	Decrease	Decrease in either/both categories	1
HMLD	HMHD	HMLD to HMHD	Increase in diversity category	3
HMLD	HMLD	No Change	No change in either category	2
HMLD	LMHD	Decrease	Decrease in either/both categories	1
HMLD	LMLD	Decrease	Decrease in either/both categories	1
HMLD	NO DATA	Decrease	Decrease in either/both categories	1
HMHD	HMHD	No Change	No change in either category	2
HMHD	HMLD	Decrease	Decrease in either/both categories	1
HMHD	LMHD	Decrease	Decrease in either/both categories	1
HMHD	LMLD	Decrease	Decrease in either/both categories	1
HMHD	NO DATA	Decrease	Decrease in either/both categories	1
NO DATA	HMHD	No Data to HMHD	Increase in migration category	4
NO DATA	HMLD	No Data to HMLD	Increase in migration category	4
NO DATA	LMHD	No Data to LMHD	Increase in diversity category	3
NO DATA	LMLD	No Data to LMHD	Increase in diversity category	3
NO DATA	NO DATA	No Change	No change in either category	2

LMLD = Low in-migration & low diversity, LMHD = Low in-migration & high diversity,
HMLD = High in-migration & low diversity, HMHD = High in-migration & high diversity

5.3.2.2 Exploratory analysis

With a number of population mixing measures quantified, the next step involved using exploratory data analysis in order to begin examining the associations between these measures and the two diseases. First, standardised incidence ratios for ALL and type 1 diabetes were calculated for CAUs categorised according to levels of population mixing, and second, the strength and directions of the relationships were determined using correlation analysis.

SIRs by population mixing categories

To begin exploring how ALL and type 1 diabetes incidence varied by different population mixing measures, SIRs were calculated for quintiles of population change and migration change

for the period 1981-2001. These two measures were chosen since population change is the most consistently used measure of population mixing adopted in the international literature to date, and is used as a surrogate for the number of new people entering an area. As argued in Chapter 2, a direct measure of the number of recent movers (e.g. migration change) is preferable. Thus a comparison of the results for each variable was thought to be of interest.

First, every CAU in New Zealand was ranked according to its population change value, and then assigned to one of five approximately equal categories. Category 1 areas were those which experienced the greatest increase in population between 1981 and 2001. Category 5 areas included areas which had decreased the most in population during the study period. Descriptive statistics for this new categorical variable were then calculated (Table 5.10). Next, the observed and expected cases of ALL at the CAU-level were summed by each population change quintile, and the corresponding SIRs, chi-square values and confidence intervals were calculated (as detailed in section 5.3.1.2). The same process was followed for the examination of type 1 diabetes incidence by population change quintile in Canterbury (Table 5.10, values in the parentheses) and for the calculations by migration change quintiles for both diseases (Table 5.11).

Table 5.10: Population change quintiles, for New Zealand
(& Canterbury in the parentheses), 1981-2001

Population change quintile	Minimum value	Maximum value	Mean value	Standard deviation	No. of CAUS
1 Highest growth	**46.97** (54.64)	**8900.00** (930.43)	**205.03** (118.22)	**575.31** (154.83)	**338** (32)
2	**19.74** (23.08)	**46.73** (53.49)	**30.70** (33.12)	**7.43** (9.41)	**338** (32)
3	**4.79** (8.49)	**19.64** (22.76)	**11.29** (14.72)	**4.16** (4.73)	**338** (32)
4	**-7.97** (1.86)	**4.78** (8.31)	**-1.50** (5.22)	**3.88** (1.95)	**338** (32)
5 Highest decline	**-100.00** (-35.76)	**-8.11** (1.71)	**-20.83** (-7.46)	**14.36** (9.49)	**338** (31)
Total New Zealand (Canterbury)	**-100.00** (-35.76)	**8,900.00** (930.43)	**44.94** (33.02)	**269.80** (82.22)	**1,690** (159)

Table 5.11: Change in the percentage of total migrants quintiles for New Zealand (& Canterbury in the parentheses), 1981-2001

Change in % total migrants quintile	Minimum value	Maximum value	Mean value	Standard deviation	No. of CAUS
1 Highest growth	**39.05** (41.95)	**355.82** (109.80)	**62.03** (56.66)	**26.68** (15.46)	**337** (32)
2	**15.78** (25.30)	**38.94** (41.38)	**27.19** (33.48)	**6.73** (4.76)	**337** (32)
3	**-13.87** (-6.44)	**15.70** (25.07)	**1.85** (13.45)	**8.87** (9.00)	**337** (32)
4	**-33.09** (-32.52)	**-14.05** (-8.10)	**-24.15** (-22.78)	**5.44** (7.09)	**337** (32)
5 Highest decline	**-72.83** (-68.85)	**-33.13** (-32.94)	**-44.44** (-45.14)	**8.00** (8.69)	**338** (31)
Total New Zealand (Canterbury)	**-72.83** (-68.85)	**355.82** (109.80)	**4.47** (7.46)	**39.95** (38.13)	**1,686** (159)

Correlation analysis

A number of correlation analyses were also carried out at the CAU-level between each population mixing variable and the raw counts of ALL and type 1 diabetes cases. Correlation analysis provides information on the strength and direction of the relationship between two variables. Thus it was employed to test whether the population mixing variables were positively or negatively associated with the two diseases, and how strong these relationships were. In addition, correlation analysis is considered an important precursor to regression analyses (Shaw and Wheeler, 1994).

In order to calculate the Pearson's Correlation Coefficient the variables must be measured on either an interval or ratio scale. In addition, this method makes the implicit assumption that the data-sample has come from a bivariate normal distribution. When this assumption is not justified, a non-parametric measure such as the Spearman's Rank Correlation Coefficient is more appropriate (Calder and Sapsford, 2006). In the case of both diseases, data were not normally distributed and therefore had to be ranked in SPSS before the Spearman's Rank Correlation Coefficient could be calculated. Regardless of the form of the correlation analysis selected, the outcome can only vary between -1.0 and +1.0, reflecting respectively a perfectively negative and perfectly positive relationship between two variables (Shaw and Wheeler, 1994).

While correlation analysis provides a useful insight into the general direction and strength of the relationships between two variables at a time, its limitations in the context of this study should be recognised. Firstly, the correlation coefficient only reflects the degree to which two variables are linearly related. Where two variables are strongly related but the relationship is non-linear, the correlation coefficient will be close to zero (Huizingh, 2007). Secondly, correlation analysis itself does not imply that one variable is dependent upon another; it merely considers the covariation of variables (Shaw and Wheeler, 1994). Finally, and critically in terms of this research, this test is bivariate in nature and very few patterns are simple enough to be accounted for satisfactorily by the variations in a single explanatory variable (King, 1969). Therefore more sophisticated analyses capable of examining the relationships with several explanatory variables, were necessary.

5.3.2.3 Regression modelling

Introduction

Thus far, the methods detailed involve the examination of univariate associations between each disease and a number of variables considered separately, including age, sex, ethnicity, year of diagnosis, area of diagnosis, and various population mixing measures. Some analyses (SIR, Poisson probability and cluster analyses) were also able to control for the population at risk. However, no single cause of ALL or type 1 diabetes has been identified, and multiple agents could operate in unison to initiate the onset of either disease (Greaves, 2002, Haverkos, 1997, Leslie and Elliott, 1994). For example, it is plausible that population mixing could work in conjunction with another risk factor to initiate higher incidence of either disease. Alternatively, any univariate relationships noted between ALL or type 1 diabetes and population mixing, may be explained by confounding risk factors. Consequently, multivariate regression analyses were employed to simultaneously assess the effects of the population mixing variables on each disease, whilst taking into account key control variables. However, traditional linear multiple regression analyses were not appropriate as both disease datasets consisted of count outcomes with low mean values (Table 5.12), and thus did not satisfy the assumption of normality. Poisson regression analysis is appropriate when the dependent variable is measured as a count, and is particularly useful where observations have very low values (Lovett and Flowerdew, 1989). As a result, Poisson regression was utilised as the starting model from which to analyse these data.

Table 5.12: Descriptive statistics of the disease observations 1980-2004

Descriptive statistics	ALL	Type 1 diabetes
Total number of observations	10,302	954
Mean value	0.0758	0.35
Variance	0.0815	0.40
Standard deviation	0.291	0.64
Minimum value	0	0
Maximum value	4	4

The first part of this section introduces the principles of Poisson regression analysis and other models employed in the analysis of count data. Next, the datasets and variables used in these analyses are described, and the modelling strategy undertaken in this research is outlined.

Poisson regression

The Poisson regression model is the most basic count regression model. It is a non-linear regression model where the probability of a count occurring is determined by the Poisson distribution (Cameron and Trivedi, 1998). In this model, the predicted value of the dependent variable for case i is the maximum likelihood estimate of the mean ($\hat{\lambda}_i$) of a Poisson distributed variable Y_i. The estimated mean $\hat{\lambda}_i$ is assumed to be logarithmically linked to a linear combination of the independent variables, giving:

$$\ln(\hat{\lambda}_i) = \beta_0 + \beta_1 x_i$$

Where β_0 is the intercept: the value of y when x is zero; and β_1 is the gradient: the amount by which y changes as x increases by one unit. This equation can also be written as:

$$\hat{\lambda}_i = \exp(\beta_0 + \beta_1 x_i)$$

This equation is the Poisson regression equivalent to the ordinary least squares regression equation (Lovett & Flowerdew 1989). The typical Poisson regression model therefore expresses the log outcome rate as a linear function of a set of predictors. These predictor variables can be composed of continuous variables, categorical variables, or some combination of both (Knudsen 1992). As with ordinary least squares regression, Poisson regression models are subject to a number of assumptions:

1. The logarithm of the dependent variable changes linearly with equal increment increases in the exposure variable.
2. Changes in the rate from combined effects of different exposures or risk factors are multiplicative.
3. Observations are independent: when an event occurs it does not affect the probability of the event occurring in the future.
4. The conditional mean of the outcome is equal to the conditional variance (equidispersion).

However, equality of the mean and variance is rarely achieved in practice. In many studies of discrete outcomes, the variance exceeds the mean value and these data can be considered to be over-dispersed (Long, 1997). Over-dispersion may result from the omission of important explanatory variables from the model (Boyle and Flowerdew, 1993). The standard model to account for over-dispersion is the negative binomial model. Negative binomial regression is a form of Poisson regression that includes a random component reflecting the uncertainty about the true rates at which events occur for individual cases (Gardner et al., 1995). In negative binomial regression, individuals are assumed to have a common mean plus a quantity specific to each individual; the random component. When all of the individuals are evaluated together, the distribution of the outcome variable is negative binomial. The usual form of the negative binomial regression model is given by:

$$\ln(\hat{\lambda}_i) = \beta_0 + \beta_1 x_i + \sigma\varepsilon$$

Where $\sigma\varepsilon$ represents the random component or error term (Byers et al., 2003). The negative binomial regression model is thus less restrictive than the Poisson regression model as it allows the variance value to exceed the mean value. It is also more conservative, and can be used to avoid overestimation of the significance of variables in models, as although the point estimates are not changed, the confidence intervals are widened (Cameron and Trivedi, 1998). Despite the potential advantages of the negative binomial model, the Poisson model has been more widely used in population mixing studies (e.g. Cox, 2007, Parslow et al., 2002, Stiller and Boyle, 1996).

A second problem is that the number of zeros in a dataset often exceeds the number predicted by either the Poisson or the negative binomial regression model (Ridout et al., 2001). As a result, zero-inflated models can be employed, which explicitly model the number of predicted zeros. For example, in zero-inflated Poisson regression, both the probability of the count outcome being

zero, and the probability of the count outcome being zero or greater following the Poisson distribution, is calculated (Afifi et al., 2007).

All of the models described above are examples of generalised linear models (GLMs) which are an extension of the linear modelling process. GLMs allow models to be fit to data that follow probability distributions other than the normal distribution. Goodness of fit for GLMs has traditionally been assessed using either the scaled deviance (twice the logarithm of the ratio of the likelihood of the data under the larger model, to that under the smaller model) or the Pearson's χ^2 statistic (sum of squares of standardised observations). However, when a high proportion of the observations are based on small values (resulting in a low mean), the approximation of the deviance by χ^2 is unreliable (Agresti, 1996, Armitage and Berry, 1994). Alternative methods of assessing model fit have been proposed, such as a simulation approach (Boyle et al., 1997) and grouping data (Wood, 2002), but are computationally complex and not widely used.

As previously noted (section 5.3.1.3), the Poisson distribution is a reasonable approximation of the ALL and type 1 diabetes datasets (Figures 5.5 and 5.6). Thus the Poisson regression model was employed as the starting model from which to analyse these data. Separate Poisson regression models were run for each disease. The count of the age group- and sex-specific cases of either ALL or type 1 diabetes in each CAU was included as the dependent variable. A number of control variables were also added as explanatory variables to these models.

Control variables

Since both childhood ALL and type 1 diabetes vary by age at diagnosis and sex, it was important to control for differences in the age and sex structure of the population at risk in each CAU. Age and sex stratified population data were obtained from Statistics New Zealand for the census years 1991, 1996 and 2001. Linear interpolations between census years were acquired from the Ministry of Health for the years 1980-2004 inclusive. For each study time period, the age group- and sex-specific population at the approximate mid-point year was employed as an estimate of the average population for each age and sex group throughout the study period. Where the approximate mid-point of the study period was a census year, census data were used as these data represent actual population counts.

It was also important to adjust for the ethnic composition of the population at risk. The incidence of both ALL and type 1 diabetes has been found to be significantly higher in children of European decent in New Zealand (Dockerty et al., 1996b, Willis et al., 2002b). While data on individual ethnicity were available for both diseases, preliminary inspection of these data revealed very small numbers of non-European cases for both ALL and type 1 diabetes, and would have resulted in a very sparse dataset. As a result, an area-level measure of the ethnic composition of each CAU was derived from census data and linear interpolations of these data. This variable was calculated as the percentage of the usually resident population whose main ethnic group was European, for the approximate mid-point of each study period.

As noted earlier in the chapter, area-level deprivation measures have been shown in some (e.g. Borugian et al., 2005, Haynes et al., 2006), but not all (Cox, 2007, Smith et al., 2006) studies to be significantly associated with childhood ALL and type 1 diabetes. In this study the New Zealand Deprivation Score 2001 (section 5.3.1.2) was employed as a proxy for social differences in hygiene, since high levels of hygiene are thought to be associated with a decreased exposure to infections in early life. In addition, neighbourhood deprivation could influence the level of population mixing in an area, since affluent areas are likely to attract more new people (Koushik et al., 2001, Stiller and Boyle, 1996). Accordingly, area-level deprivation could confound the relationship between population mixing and the diseases, and thus needed to be controlled for in the analyses.

Population density was used as a proxy measure for the intensity of contacts between individuals. In high population density areas people are more likely to come into contact with a higher number of other people (Parslow et al., 2001). Population density has been found to be significantly related to ALL and type 1 diabetes in several previous studies (e.g. Feltbower et al., 2005, Hjalmars and Gustafsson, 1999, Karvonen et al., 1997b). This variable was calculated for the years 1991 and 2001 by dividing the total population of each CAU in New Zealand by its area (km^2).

Measures of household overcrowding were also created to account for the intensity of social contacts, but this time at the household/family-level. Household overcrowding has been linked to the spread of infectious diseases in children (Smith et al., 1990) and is thus thought to be protective against the later development of childhood type 1 diabetes (e.g. Patterson et al., 1996, Staines et al., 1997). A similar argument can be made for childhood ALL. No data on household/family size were available for individual children therefore an area-level measure was

created. Using census data for 1991, 1996 and 2001 at the CAU-level, the number of households with six or more usual members was calculated as a percentage of the total number of households in each area. Unfortunately no information on the number of people per room was available. Areas with a high percentage of households with six or more usual members were considered to be a reasonable proxy for areas with a high level of household overcrowding. While overcrowding has been used as a substitute measure for the protective effect of early life infectious exposure, it is also useful to examine whether smaller households (particularly with only one child) are associated with a higher incidence of either disease. Thus the percentage of households with three or less usual members was also calculated. It was postulated that a child with no siblings (living in a three person household or less) would have fewer infections in early life compared to children with siblings.

Finally, the Statistics New Zealand urban/rural classification of each CAU in 2001 (section 5.3.1.2) was included to control for the urban/rural nature of areas across New Zealand. A summary of all of the control variables used in the analyses is given in Table 5.13.

Table 5.13: Details of the control variables used in the regression analyses

Control variables	Data sources	Years
Individual-level (for each case)		
Age group at diagnosis (0-4, 5-9, 10-14 yrs) Sex (male, female)	National Cancer Registry Canterbury Diabetes Register	For every year in each study period
Area-level (for each CAU)		
Age- and sex-specific population at risk (males: 0-4, 5-9, 10-14, and females: 0-4, 5-9, 10-14)	Census counts (Statistics New Zealand) and linear interpolations of census counts (Ministry of Health)	Approximate mid-point of each study period
% European	Census counts (Statistics New Zealand) and linear interpolations of census counts (Ministry of Health)	Approximate mid-point of each study period
Deprivation score	New Zealand Deprivation Score (Salmond & Crampton 2002)	2001
Population density	Calculated using census counts & CAU area calculations in a GIS	1991, 2001
% Households < 3 people	Census counts (Statistics New Zealand)	1991, 1996, 2001
% Households > 6 people	Census counts (Statistics New Zealand)	1991, 1996, 2001
Urban/rural category	Statistics New Zealand Urban/rural Profile	2001

Population mixing variables

The population mixing aims of this thesis were to:

- Test a modification of Kinlen's (1988) theory that *increases* in population mixing could act as a trigger for ALL or type 1 diabetes in children.
- Test the hygiene hypothesis that low population mixing levels in early life are associated with an increased risk of childhood type 1 diabetes.

Consequently the population mixing change measures listed in Table 5.14 were used in both the ALL and type 1 diabetes analyses. These measures were used to examine whether a change from low to high population mixing levels was associated with higher incidence of either ALL or type 1 diabetes (triggering effect). The static population mixing measures listed in Table 5.14 were used only in the type 1 diabetes analyses. These static measures were ascertained at the beginning of each study period to test whether high levels of population mixing were associated with low type 1 diabetes incidence (protective effect)

Table 5.14: Final population mixing measures used in the ALL and type 1 diabetes regression analyses

Population mixing measure	ALL	Type 1 diabetes
Population mixing change measures		
Population change	✓	✓
Change in % of total migrants	✓	✓
Change in % of child migrants	✓	✓
Change in % of overseas migrants	✓	✓
Change in % of overseas visitors	✓	✓
Change in one year mobility %	✓	✓
Change in migrant diversity	✓	✓
Population mixing change category:	✓	✓
1. Decrease in either/both categories	✓	✓
2. No change in either category	✓	✓
3. Increase in diversity category	✓	✓
4. Increase in migration category	✓	✓
Static population mixing measures		
% of total migrants		✓
% of child migrants		✓
% of overseas migrants		✓
% of overseas visitors		✓
One year mobility %		✓
Migrant diversity		✓
Population mixing category:		✓
1. Low in-migration & low diversity		✓
2. Low in-migration & high diversity		✓
3. High in-migration & low diversity		✓
4. High in-migration & high diversity		✓

The effects of population mixing were examined over a number of different timescales since it is unclear over what length of time population mixing might be important for disease pathogenesis in children. Shorter time scales are likely to be more accurate in capturing static population mixing levels or changes in population mixing appropriate to a child's actual exposure. In addition, longer time scales may not capture important population movements. However, longer time periods are able to take into account longer lag times between population mixing exposures, disease initiation and the subsequent disease diagnosis. For example, for children who were aged 14 years at disease diagnosis, early life exposure to infections through population mixing would have occurred 12/13 years prior to the diagnosis. Consequently, the effects of population mixing on ALL and type 1 diabetes were examined for the 25 year study period as a whole (1980-2004),

two time periods of 12 and 13 years (1980-1991 and 1992-2004), and four time periods of six/seven years (1980-1986, 1987-1992, 1993-1998 and 1999-2004) (Table 5.15).

Table 5.15: Diagnosis periods and associated population mixing measurements

Diagnosis period	Number of years	Population mixing change measure	Static Population mixing measure
1980-2004	25	1981-2001	N/A
1980-1991	12	1981-1991	1981
1992-2004	13	1991-2001	1991
1980-1986	7	1981-1986	1981
1987-1992	6	1986-1991	1986
1993-1998	6	1991-1996	1991
1999-2004	6	1996-2001	1996

The timing of the population mixing measurements (static and change measures) was constrained by the use of census data which is only collected every five years in New Zealand. The population mixing change measures aimed to cover roughly the same time periods as the diagnosis periods. However, for the 6/7 year analyses the population mixing change measures were staggered slightly, with the later three periods (1987-1992, 1993-1998 and 1999-2004) incorporating a lag time between the beginning of the population mixing change period, and the beginning of the diagnosis period (1, 2 and 3 years respectively) (Table 5.16). For childhood leukaemias in general, estimates of the lag time between disease initiation and the appearance of diagnosable symptoms vary from between 3-12 months (Greaves, 2006) to 2-4 years (Langford, 1991). For childhood diabetes, disease initiation could occur 'months to years' before recognisable signs of the illness begin to show (Petrovsky and Schatz, 2003). Thus it was considered appropriate to use measures which included this range of different lag times.

The static population mixing measures were determined at the beginning of each diagnosis period in order to approximate early life infectious exposure for the majority of children. In the 6/7 year analyses, varying lag times were again incorporated into the last three diagnosis periods and it was considered of interest whether these would affect the results (Table 5.16). No 25 year period was examined for the static population mixing measures, since a measurement at the beginning of the study period (1981) was unlikely to be appropriate for childhood diabetes cases diagnosed in the late 1990s and early 2000s.

Table 5.16: Timing of the population mixing (PM) measurements

Diagnosis period	Population mixing change analyses		Static population mixing analyses	
	PM Measurement	***Difference between start PM year and start diagnosis year***	***PM Measurement***	***Difference between PM year and start diagnosis year***
1980-1986	1981-1986	-1	1981	-1
1987-1992	1986-1991	+1	1986	+1
1993-1998	1991-1996	+2	1991	+2
1999-2004	1996-2001	+3	1996	+3

Analyses

The ALL analyses were conducted for a total of 1,717 out of a possible 1,859 CAUs in New Zealand. CAUs were removed from the analysis if their population in 1991 was zero, and their description implied that they represented a water body or small island. Where the 1991 population was zero but the CAU description did not relate to a water body/small island, the population was changed to 1 and these CAUs were retained in the analyses.

For each CAU, the population and ALL counts were calculated for six age-sex groups (Table 5.17). Therefore, each CAU was represented six times in the dataset, which therefore included a total of 10,302 observations (6 x 1,717 CAUs). The age- and sex-specific population at risk for each CAU (measured at the approximate mid-point of the study period) was included as an exposure variable in the models. Two categorical variables were included to define which age and sex group each record represented. The values for the control (Table 5.13) and population mixing change (Table 5.14) variables were also associated with each record. The area-level variables (apart from the population at risk) were not age- and sex-specific, and were thus repeated six times within the dataset. The descriptive statistics for all of the variables included in the ALL analyses are given in Tables 5.17-5.19.

Table 5.17: Descriptive statistics for the continuous variables used in the ALL analysis (1980-2004)

Variable	Minimum	Maximum	Mean	Standard deviation
ALL counts:				
Male 0-4 years	0	4	0.15	0.41
Female 0-4 years	0	4	0.11	0.34
Male 5-9 years	0	3	0.07	0.27
Female 5-9 years	0	2	0.05	0.23
Male 10-14 years	0	2	0.05	0.22
Female 10-14 years	0	2	0.03	0.18
Control variables				
Age- & sex-specific population	1	501	76.08	58.23
% European	3.36	100.00	79.82	17.88
Deprivation score	858	1322	995.32	80.77
Population density	0.00	5,886.99	938.91	1,006.69
% Households < 3 people	0.00	100.00	68.55	10.48
% Households > 6 people	0.00	50.00	4.93	4.21
Population mixing change variables				
Population change	-100.00	8900	44.38	267.61
Change in % total migrants	-72.83	355.82	4.38	39.59
Change in % child migrants	-100.00	432.92	-6.91	44.99
Change in % overseas migrants	-100.00	1,957.85	89.48	124.86
Change in % overseas visitors	-98.40	30,108.33	358.59	1,223.81
Change in migrant diversity	-60.45	1,811.38	7.70	60.63

Table 5.18: ALL cases and number of CAUs by urban/rural categories (1980-2004)

Urban/rural category	Number of CAUs	Number of ALL cases
Main urban area	6,018	546
Satellite urban community	276	27
Independent urban community	1,248	95
Rural area with high urban influence	432	25
Rural area with moderate urban influence	576	24
Rural area with low urban influence	1,260	52
Highly rural/remote area	474	12
None assigned	18	0

Table 5.19: ALL cases and number of CAUs by population mixing change categories (1980-2004)

Population mixing change category	Number of CAUs	Number of ALL cases
1. Decrease in either/both categories	3,654	228
2. No change in either category	3,444	250
3. Increase in diversity category	1,032	84
4. Increase in migration category	2,172	219

A total of 159 out of a possible 166 CAUs located in the Canterbury study area were included in the type 1 diabetes analyses. Seven CAUs were removed as they represented lakes or estuaries. Since each CAU was represented six times in the dataset (once for every age-sex group), a total of 954 records were analysed. For each CAU the number of cases of type 1 diabetes as well as the count of the population was calculated for six age-sex groups (Table 5.20). Two categorical variables were included to define which of the age-sex groups each record belonged to. The population at the approximate mid-point of each study period for each age and sex group was included as an exposure variable in the models. All of the other explanatory variables (population mixing and the control variables) were not age- and sex-specific and hence the estimate was identical for the six times each CAU was included in the dataset. The descriptive statistics for the type 1 diabetes variables are given in Tables 5.20-5.23.

Table 5.20**:** Descriptive statistics for the continuous variables used in the type 1 diabetes analysis (1980-2004)

Variable	Minimum	Maximum	Mean	Standard deviation
Type 1 diabetes counts:				
Male 0-4 years	0	3	0.23	0.53
Female 0-4 years	0	2	0.26	0.48
Male 5-9 years	0	2	0.33	0.57
Female 5-9 years	0	3	0.31	0.58
Male 10-14 years	0	3	0.47	0.69
Female 10-14 years	0	4	0.52	0.84
Control variables				
Age- & sex-specific population	1	282	75.24	47.41
% European	59.71	99.78	90.96	5.68
Deprivation score	878	1170	966.28	58.29
Population density	0.36	3,957.17	1,198.75	1,083.44
% Households < 3 people	45	92.03	71.14	8.42
% Households > 6 people	0	7.38	3.12	1.47
Static population mixing variables				
% Total migrants	24.16	96.30	53.57	19.26
% Child migrants	2.34	23.50	10.28	5.05
% Overseas migrants	0.00	12.47	2.80	1.80
% Overseas visitors	0.00	42.01	1.08	3.66
Migrant diversity	0.80	3.54	1.80	0.52
Population mixing change variables				
Population change	-35.76	930.43	33.02	82.22
Change in % total migrants	-68.85	109.80	7.46	38.13
Change in % child migrants	-72.31	99.79	-6.38	36.71
Change in % overseas migrants	-78.28	421.85	105.43	92.08
Change in % overseas visitors	-62.26	2,930.30	285.01	380.49
Change in migrant diversity	-37.01	254.93	10.67	36.19

Table 5.21: Type 1 diabetes cases and number of CAUs by urban/rural categories (1980-2004)

Urban/rural category	Number of CAUs	Number of type 1 diabetes cases
Main urban area	690	273
Satellite urban community	60	29
Independent urban community	6	0
Rural area with high urban influence	48	15
Rural area with moderate urban influence	102	12
Rural area with low urban influence	6	0
Highly rural/remote area	42	8

Table 5.22: Type 1 diabetes cases and number of CAUs by population mixing categories (1980-2004)

Population mixing category	Number of CAUs	Number of type 1 diabetes cases
1. Low in-migration & low diversity	282	121
2. Low in-migration & high diversity	252	97
3. High in-migration & low diversity	354	107
4. High in-migration & high diversity	66	12

Table 5.23: Type 1 diabetes cases and number of CAUs by population mixing change categories (1980-2004)

Population mixing change category	Number of CAUs	Number of type 1 diabetes cases
1. Decrease in either/both categories	318	103
2. No change in either category	342	135
3. Increase in diversity category	54	12
4. Increase in migration category	240	87

Modelling strategy

A number of strategies exist for examining multivariate relationships, all of which depend upon the type of data and the objectives of the research (Shaw and Wheeler, 1994). This study aimed to test whether population mixing was significantly associated with either ALL or type 1 diabetes after controlling for possible confounding variables. The first step of the modelling strategy involved choosing an adequate model to represent these data.

ALL analysis 1980-2004

Since the ALL variance value was greater than the mean value (Table 5.12), these data can be considered over-dispersed. Consequently Poisson and negative binomial regression models were statistically compared in STATA to check which model better fit these data. Negative binomial regression models were fitted to a random sample of models, and a likelihood ratio test for over-dispersion was performed. In this test, when alpha is equal to zero, the negative binomial and Poisson models are ultimately the same, so the Poisson model can still be considered appropriate (Long, 1997). For all four of the randomly picked models (Table 5.24) the p-value of the test statistic (that alpha is equal to zero) was less than 0.05 and thus considered significant. This finding showed that in each of these models, alpha was significantly different to zero and therefore that the negative binomial regression model was a better fit to these data.

Since 93.03 percent of the observations had an ALL count of zero (Table 5.25), it was also necessary to test if a zero-inflated negative binomial model was more appropriate than the standard negative binomial model. Accordingly, zero-inflated negative binomial regression models were run for the same sample of models as before, together with the Vuong test to statistically compare the zero-inflated and standard models. The Vuong test statistic (z) has a standard normal distribution with large positive values favouring the zero-inflated model and with large negative values favouring the non zero-inflated version. Values close to zero are in absolute favour neither model (Long, 1997, Vuong, 1989). According to their p-values (Table 5.26), none of the z-statistics were significantly large enough to support the zero-inflated model, and therefore the standard negative binomial model was used for all of the subsequent ALL analyses for the 1980-2004 period. The same strategy for choosing an appropriate model was adopted for every year group in the study, and the final chosen models for the ALL analyses are listed in Table 5.27.

Table 5.24: Likelihood ratio test results for negative binomial regression models of ALL counts and various explanatory variables 1980-2004

Variable included	Alpha	Test Statistic	P-value
Age & sex	0.370	6.61	0.005
Change in overseas migrants	0.731	18.20	0.000
Age, sex & population mixing change category	0.355	6.16	0.007
Age, sex, population density & population change	0.326	5.44	0.010

Table 5.25: ALL counts 1980-2004 by frequency and percentage

ALL count	Frequency	Percentage
0	9,584	93.03
1	664	6.45
2	47	0.46
3	5	0.05
4	2	0.02

Table 5.26: Vuong test results for zero-inflated negative binomial regression models of ALL count and various explanatory variables, 1980-2004

Variable included	Z-Statistic	P-value
Age & sex	0.54	0.293
Change in overseas migrants	0.12	0.454
Age, sex & population mixing change category	1.39	0.082
Age, sex, % European & population change	0.97	0.167

Table 5.27: Best fitting regression models for the ALL data by study period

Study period	Number of years	Total ALL cases	Best fitting regression model
1980-2004	25	781	Negative binomial
1980-1991	12	342	Poisson
1992-2004	13	439	Negative binomial
1980-1986	7	183	Poisson
1987-1992	6	185	Zero-inflated Poisson
1993-1998	6	194	Zero-inflated Poisson
1999-2004	6	219	Poisson

Type 1 diabetes analysis 1980-2004

The type 1 diabetes data for the Canterbury region (1980-2004) displayed similar characteristics to the ALL cases for the whole of New Zealand. Again, the variance (0.40) was greater than the mean value (0.35) (Table 5.12) and thus it was important to test whether these data were significantly over-dispersed. As a result, the fit of Poisson and negative binomial models were compared using the likelihood ratio test for over-dispersion in a random sample of models. The p-value of the test statistic (that alpha is equal to zero) in all four of the models was greater than 0.05 and thus was not considered significant (Table 5.28). Consequently, there was no significant difference between the negative binomial and Poisson models; therefore the Poisson model was an adequate fit to these data.

As with the ALL data, the majority (71.91 percent) of observations had a count of zero type 1 diabetes cases during the study period (Table 5.29). As a result it was necessary to test whether a zero-inflated Poisson regression model was a better fit than the standard Poisson regression model. The Vuong test statistics and their p-values (Table 5.30) revealed that none of the z-statistics were significantly large enough to support the zero-inflated model, and therefore the standard Poisson model was used in all of the subsequent type 1 diabetes analyses for this time period. The final models chosen for the other study periods are listed in Table 5.31.

Table 5.28: Likelihood ratio test results for negative binomial regression models of type 1 diabetes counts and various explanatory variables, 1980-2004

Variable included	Alpha	Test Statistic	P-value
Age & sex	5.30e-06	0.0e+00	0.500
Change in overseas migrants	0.0587541	0.25	0.308
Age, sex & population mixing change category	4.66e-06	0.0e+00	0.500
Age, sex, population density & population change	6.20e-08	0.0e+00	0.500

Table 5.29: Type 1 diabetes counts 1980-2004 by frequency and percentage

Type 1 Diabetes Count	Frequency	Percentage
0	686	71.91
1	211	22.12
2	47	4.93
3	8	0.84
4	2	0.21

Table 5.30: Vuong test results for zero-inflated Poisson regression models of type 1 diabetes count and various explanatory variables, 1980-2004

Variable included	Z-Statistic	P-value
Age & sex	0.91	0.1818
Change in overseas migrants	0.35	0.3634
Age, sex & population mixing change category	0.86	0.1952
Age, sex, % European & population change	-2.76	0.9971

Table 5.31: Best fitting regression models for the type 1 diabetes data by study period

Study period	Number of years	Total type 1 diabetes cases	Best fitting regression model
1980-2004	25	337	Poisson
1980-1991	12	111	Poisson
1992-2004	13	226	Poisson
1980-1986	7	55	Zero-inflated Poisson
1987-1992	6	72	Zero-inflated Poisson
1993-1998	6	90	Poisson
1999-2004	6	120	Poisson

The next stage involved checking whether any of the explanatory variables (Tables 5.13 and 5.14) needed to be transformed (Lovett et al., 1986). This step was carried out on data for the study period as a whole (1980-2004) for each disease, as an indicator for other study periods. All of the explanatory variables for both diseases were considered to be significantly different to the normal distribution according to their Kolmogorov-Smirnov test statistics (Field, 2005). Variables that exhibited a positive skew were therefore logged and those which were negatively skewed were squared. The coefficients, standard errors and p-values of these transformed variables were compared in univariate regression models to those of the raw variables. No significant differences were noted in any of the results, for either disease. Consequently, untransformed variables were utilised in all of the subsequent models for ease of interpretation.

The next stage involved examining the univariate relationships between each disease and each of the explanatory variables. First, models containing only an intercept (null model) were fitted in order to provide a measure of the variation in the dependent variable around its mean. The deviance value from the null model was used as a baseline measure of fit against which subsequent models were compared. Second, all of the explanatory variables (Tables 5.13 and 5.14) were then modelled separately to examine the univariate relationships between each disease and each control and population mixing variable. The age- and sex-specific population at the mid-point of the study period was included as an exposure variable in each model to adjust for differences in the population at risk between areas. Robust standard errors were calculated to account for within area clustering, since values for the area-level variables were repeated six times in each dataset.

Two of the continuous population mixing variables (change in total migrants and migrant diversity) were also analysed as categorical variables for the 1980-2004 analyses to investigate

the possibility of a threshold effect (e.g. Parslow et al., 2001, Stiller and Boyle, 1996). Categorisation also removes the assumption of linearity in the relationship between log incidence and population mixing (Parslow et al., 2001). However, no significant trends were noted by categories, therefore these variables were kept in their continuous form. Furthermore, the raw numbers of total and child migrants were also tested in the 1980-2004 analyses for both diseases. The associations found were in a similar direction to the percentage total and child migrant variables, but were much weaker. Consequently these variables were not included in further analyses.

Finally, a number of multivariate regression models were formulated with the aim of testing whether any associations found between either disease and the population mixing variables remained after controlling for potential confounders. Many different methods exist for building multivariate regression models. For example there are several systematic and mechanical algorithms which can be employed to find suitable multivariate models, such as stepwise and best-subset regression. However, it is becoming increasingly acknowledged that inclusion of variables should be more firmly based on theoretical rather than statistical measures (Sterne et al., 2001, Tilling et al., 2005). In order to specify credible models, knowledge of the background scientific literature is essential. This knowledge should include information about relations of potential confounders to the study exposures and study diseases, as well as relations of study exposures to the study diseases (Rothman and Greenland, 1998). As a result, this research used theory guided by the literature and the results of the univariate control models for the 12/13 and 25 year analyses, to formulate a number of multivariate models for each study period. Each population mixing measure was then added separately to each of these models (Figure 5.7). Details regarding the formulation of the multivariate models and their results are given in the two results chapters for each disease (Chapters 7 and 8). The next chapter (6) of this thesis considers patterns of population mixing for small areas across New Zealand in more detail.

Summary of the modelling strategy:

1. Initially, the appropriate type of probability distribution was chosen to model the data. Examination of the descriptive statistics revealed that the Poisson distribution was an appropriate starting model. Thus, at the outset, Poisson models were fitted to random combinations of variables. These models were then statistically compared with negative binomial and zero-inflated models to determine which model better fitted the data.
2. After the appropriate type of regression model was selected, each explanatory variable was examined to check whether they should be transformed for inclusion in the models.
3. In the first instance, models containing only an intercept were fitted in order to provide a measure of the variation in the dependent variable around its mean.
4. Next all of the explanatory variables were then modelled separately. The age and sex-specific population at the mid-point of the study period was included as an exposure variable in each model and robust standard errors were calculated to account for within area clustering.
5. Using theory guided by the literature and the results of the univariate control models for the 12/13 and 25 year analyses, a base model and alternative multivariate models were constructed. Where univariate results for the year group being analysed indicated additional significant control variables, the base and alternative models were altered accordingly.
6. Each population mixing measure was then added separately to each of these models.

Figure 5.7: Summary of the regression modelling strategy employed

Chapter 6: Population mixing in New Zealand

6.1 Introduction

As detailed in the previous chapter, a number of different population mixing measures were created for this research in order to capture a range of population movements in New Zealand during the study period. This chapter examines a selection of these measures in detail to determine how they vary both temporally and spatially around the country. It concentrates on measures of population mixing *change*, since these measures are used in the analyses for both diseases. Spatial patterns of these measures are generally examined for two ten year periods (1981-1991 and 1991-2001) and temporal and contextual patterns are examined for four five year periods (1981-1986, 1986-1991, 1991-1996, and 1996-2001).

6.2 Total population change

Many population mixing studies have simply used measures of population growth as a proxy for the number of new people entering an area. Population growth has been steadily occurring in New Zealand since at least 1858 when the country held its first national (including both Māori and non-Māori counts) census. In 1858 the total population was 115,461 compared to 3,129,384 in 1976 (Statistics New Zealand, 2007a).

6.2.1 Description of population change 1981-2001

The national population of New Zealand continued to grow during the study period of this research. The total New Zealand population increased from 3,142,800 in 1981 to 3,736,269 in 2001, an increase of 593,469 people or almost 20 percent (Table 6.1). The largest five year increase occurred between 1991 and 1996 when the population grew by approximately seven percent. For other five year periods the increase was around three to four percent. Considering the two ten year periods, the largest population growth (10.75 percent) was noted between 1991 and 2001 (Table 6.1).

Table 6.1: Population change in New Zealand between 1981 and 2001

Time Period	Population Change (%)	Population Change (numbers)
1981-1986	3.79	118,986
1986-1991	3.43	111,852
1991-1996	7.21	243,240
1996-2001	3.30	119,391
1981-1991	7.34	230,838
1991-2001	10.75	362,631
1981-2001	18.88	593,469

Despite population increases at the national-level (Table 6.1), not all of New Zealand's 74 territorial authorities (TAs) witnessed a net population growth during the study period (Table 6.2 and Figure 6.1). Considering five year time intervals, the minimum population change values ranged from -6.43 percent between 1991 and 1996, to -42.49 percent between 1981 and 1986 (Table 6.2). Thus the population of some TAs almost halved during the earliest time period. The maximum growth values also varied over time, with the highest growth noted during the period 1991-1996 (41.80 percent) and the lowest during 1996-2001 (19.81 percent). The average growth of TAs was less than five percent for all of the five year time spans, and was highest between 1991 and 1996 (4.97 percent). The largest variation in population change between TAs occurred between 1981 and 1986 (standard deviation 8.38) (Table 6.2).

Examining population change for two ten year periods (1981-1991 and 1991-2001), revealed slightly higher maximum (up to 69.90 percent), minimum values (up to -48.98 percent) and average values (up to 5.75 percent), compared to the five year time periods (Table 6.2). A number of TAs in the South Island, including some in the West Coast, Southland, Otago and Canterbury regions, decreased in population during both decades (Figure 6.1). The largest decrease in population (48.98 percent) was noted in the Mackenzie District in the Canterbury region between 1981 and 1991. For the period 1991-2001, the largest decrease in population (16.61 percent) occurred in the Kawerau District in the Bay of Plenty in the North Island. TAs which decreased in population in the North Island tended to cluster in the west, in the regions of Taranaki, Manawatu-Wanganui and south-west Waikato, especially in the latter half of the study period.

The average population growth for New Zealand TAs was just over five percent during both 1981-1991 and 1991-2001 (Table 6.2). Population growth at the TA-level was generally concentrated in the north of the North Island and north-east of the South Island for both time periods (Figure 6.1). However, the greatest increase in population (60.52 percent) between 1981 and 1991 occurred in the Queenstown-Lakes TA, in the southern half of the South Island. Moreover, during 1991-2001, the population in this TA increased by a further 69.90 percent. Other large population increases were

noted in the Rodney District just north of Auckland (48.69 percent during 1981-1991 and 39.17 during 1991-2001) and the Kapiti Coast District, north of Wellington (32.46 percent during 1981-1991 and 21.49 percent during 1991-2001). The area covered by the Canterbury Diabetes Register, witnessed an increase in population in both decades, with the largest growth (32.06 percent) occurring in the Waimakariri District between 1991 and 2001 (Figure 6.1).

Table 6.2: Descriptive statistics of population change
in New Zealand territorial authorities between 1981 and 2001

Time Period	**Relative Population Change (%)**				
	Minimum	**Maximum**	**Mean**	**Median**	**Standard Deviation**
1981-1986	-42.49	33.81	3.52	2.56	8.38
1986-1991	-12.50	21.19	1.81	0.94	6.92
1991-1996	-6.43	41.80	4.97	3.23	7.60
1996-2001	-14.39	19.81	0.26	-0.64	6.60
1981-1991	-48.98	60.52	5.75	4.18	14.40
1991-2001	-16.61	69.90	5.66	1.74	14.51
1981-2001	-47.90	172.72	13.37	6.73	31.93

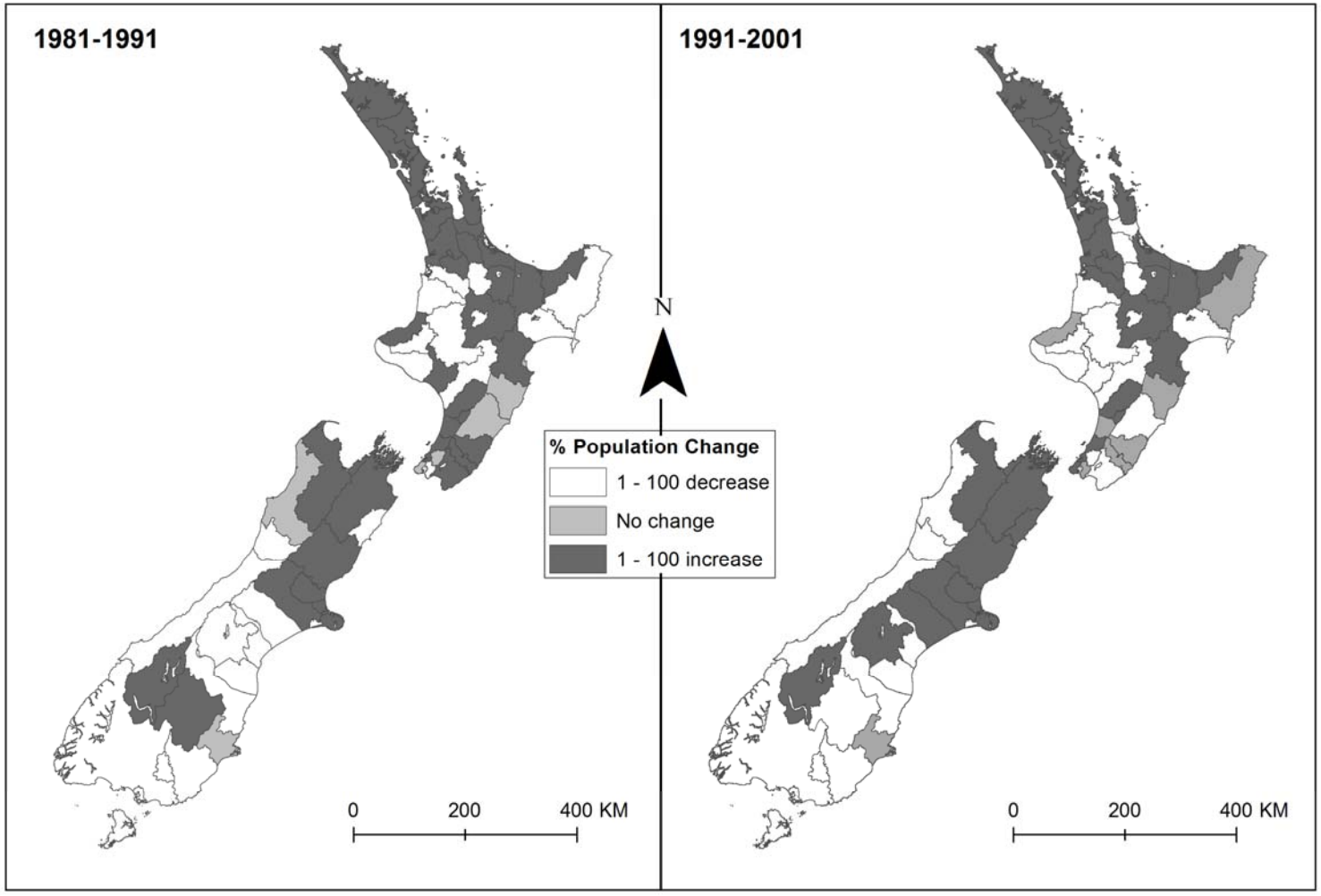

Figure 6.1: Relative population change by territorial authority area 1981-1991 and 1991-2001

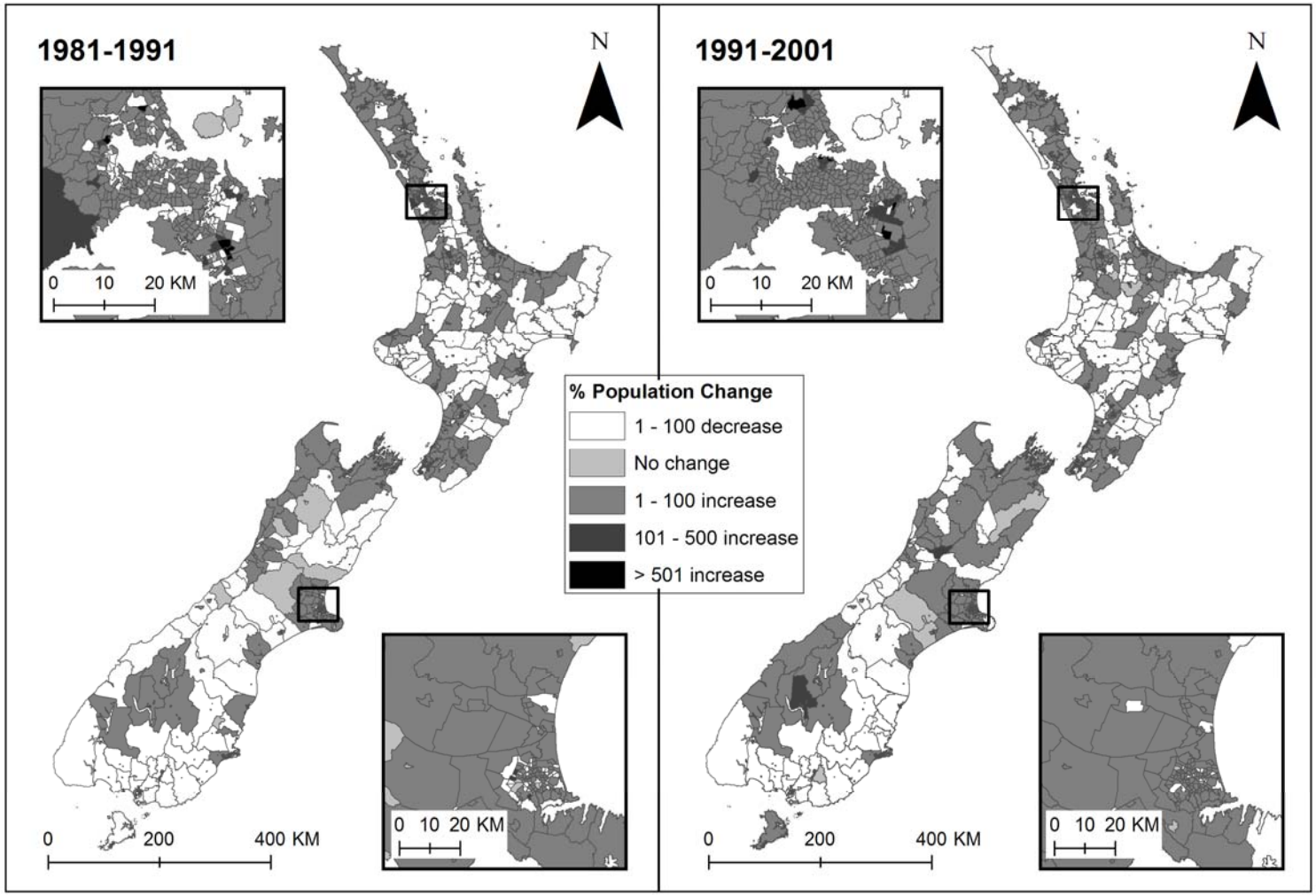

Figure 6.2: Relative population change by census area unit (CAU) with insets of Auckland in the North Island and Christchurch in the South Island 1981-1991 and 1991-2001

Using census area unit (CAU) boundaries that were consistent over time, it was possible to directly compare changes in the population composition at a finer geographical scale (Table 6.3 and Figure 6.2). The highest average population growth at the CAU-level was 9.59 percent and occurred between 1991 and 1996. This figure is more than double the growth noted for the previous period (1986-1991). The largest maximum growth of a single CAU was 3,613.33 percent and occurred in southern Auckland between 1996 and 2001. Interestingly, a number of CAUs in each time period lost their total population. These areas tended to have a small absolute population at the beginning of the time period and were often port/harbour areas. Figures for the modal value for each period are also reported due to the skewed distribution of the data. However, not much variation is noted between the different time periods for this measure.

Table 6.3: Descriptive statistics of population change in New Zealand CAUs between 1981 and 2001

Time Period	Relative Population Change (%)					
	Minimum	Maximum	Mean	Mode	Median	Standard Deviation
1981-1986	-100.00	600.00	7.69	2	2.25	29.15
1986-1991	-100.00	410.53	4.45	1	1.61	25.02
1991-1996	-100.00	432.35	9.59	2	5.31	27.86
1996-2001	-100.00	3613.33	7.77	1	0.72	99.97
1981-1991	-100.00	2152.94	15.79	-5	4.10	74.38
1991-2001	-100.00	2314.71	20.23	5	5.48	114.31
1981-2001	-100.00	8900.00	44.94	5	10.88	269.79

Considering the geographical patterns at the CAU-level for the ten year period between 1981 and 1991 (Figure 6.2), a number of trends are apparent. In the North Island, population growth was concentrated around the main towns and cities, with areas of decline generally occurring in the more rural areas in the centre of the island. The northern and southern tips of the island, and the Bay of Plenty region in the north-east, generally saw a rise in population during this decade. In the South Island, large areas of population loss were also noted in the more rural CAUs, especially on the south-west coast, in the far south, and in the mountainous Mackenzie District. Areas which increased in population included the north-west, and north coast of the island, parts of the main population centres of Christchurch, Dunedin and Invercargill, and also some CAUs in the Central Otago, Queenstown-Lakes, and Southland TAs.

Turning to the second decade (1991-2001) of the study period (Figure 6.2), the spatial patterns of population growth and decline in the North Island were generally consistent with the earlier period. Again, CAUs in the far north and south of the island and the Bay of Plenty area, generally noted population increases, with some exceptions in the Far North, Kaipara, and Upper and Lower Hutt TAs. However, unlike the earlier period, almost all of the CAUs in Auckland city and surrounding

areas increased in population between 1991 and 2001. Large areas of population growth also occurred in the south west of the South Island during this period, reaching parts of Westland on the west coast. Increases also appeared in larger parts of the Canterbury region, especially within the Selwyn TA, and in the majority of the north part of the South Island, but with some exceptions on the West Coast. General population decline occurred in the south east of the South Island, with exceptions around the outskirts of Oamaru, Dunedin, Gore and Invercargill, where population increases were recorded.

6.2.2 Contextual understanding of population change 1981-2001

As well as key spatial differences in population change over the study period, there were also important variations by deprivation and urban/rural status. When CAUs were divided into deprivation quintiles, the highest relative growth in population was experienced in the most affluent fifth of CAUs in New Zealand for all of the five year time periods (Figure 6.3). In general, relative population growth decreased with increasing deprivation, and the least growth occurred in the most deprived areas for every time period. No overall decline in population was noted in any of the deprivation quintiles, for any period. Considering trends in absolute population growth over time, the most growth in every deprivation quintile occurred between 1991 and 1996 (Figure 6.4). In the least deprived fifth of New Zealand CAUs, absolute population growth increased steadily from 28,950 in 1981-1986 to 65,004 in 1991-1996, but then decreased to 48,939 for the period 1996-2001. In the most deprived areas of New Zealand, population growth decreased from 20,082 in 1981-1986 to 15,282 in 1986-1991, increased to a high of 32,958 in 1991-1996, and then decreased to a low of 1,215 by 1996-2001. The five year period to exhibit the most variation by deprivation quintile was the most recent period, 1996-2001. During this time, the ratio of absolute population change from quintile 1 to 5 was 40.3, compared to only 1.4 for the earliest period, 1980-1986. Polarisation of population growth by deprivation category in New Zealand has therefore increased over time.

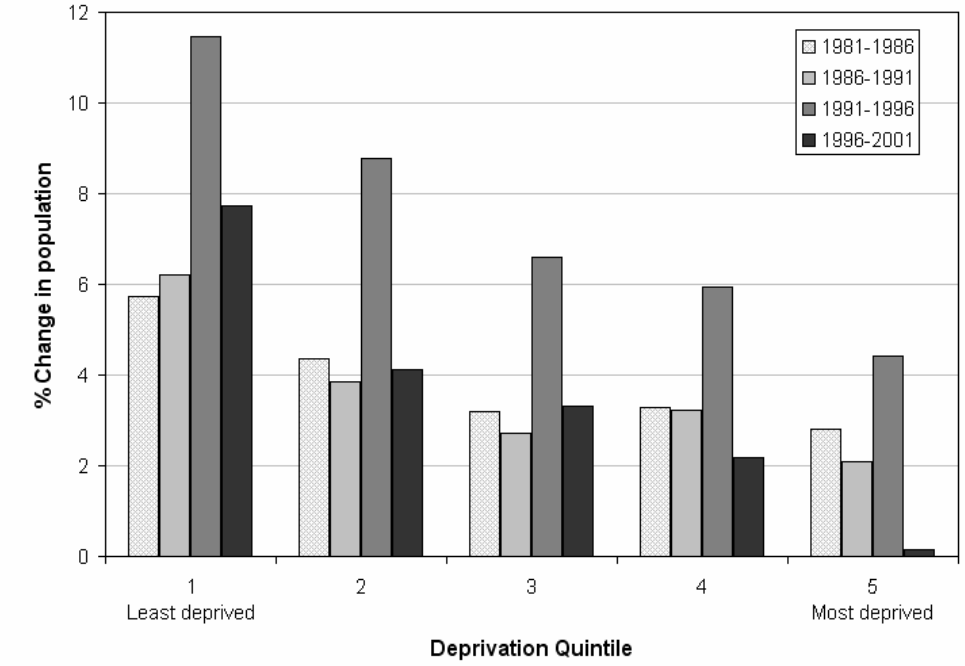

Figure 6.3: Relative population change by deprivation quintile 1981-2001

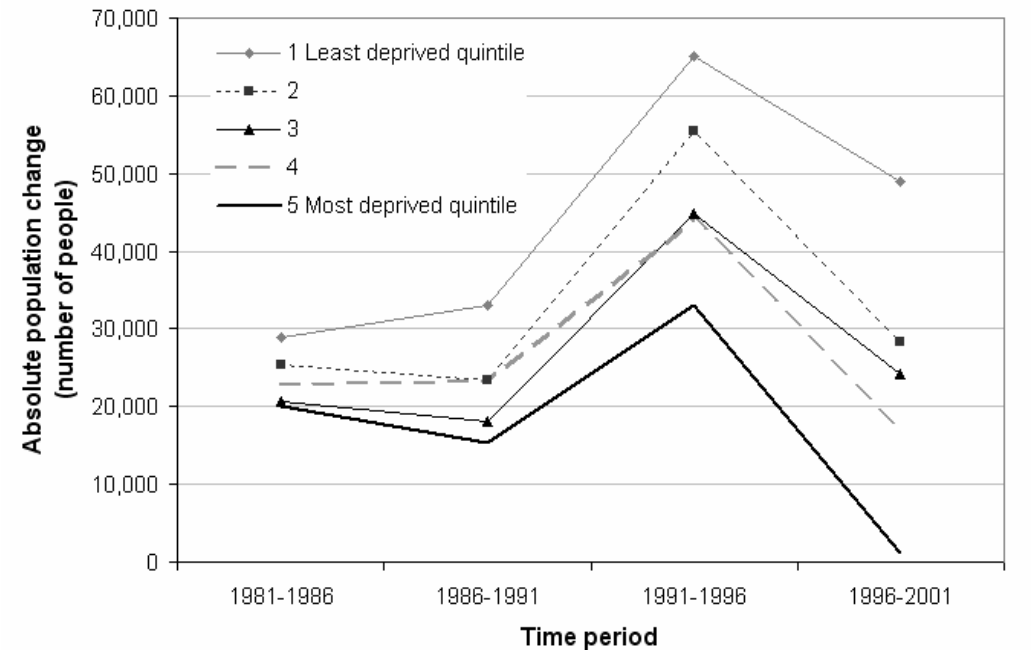

Figure 6.4: Absolute population change by deprivation quintile 1981-2001

In terms of urban/rural trends, the highest relative growth for all four of the five year time spans was witnessed in rural areas with a high urban influence (Figure 6.5). In these areas, population growth decreased slightly from 12 percent in the period 1981-1986 to 11 percent in 1986-1991, and then increased to almost 18 percent by 1991-1996, before decreasing to a low of nine percent in 1996-2001. Main urban areas, satellite urban communities and rural areas with a moderate urban influence

also increased in population during every time period, with the largest increases generally occurring between 1991 and 1996. The population of independent urban communities grew slightly between 1981 and 1996 but decreased in the final five years of the study. The change in population in rural areas with a low urban influence fluctuated between increase and decline over alternate time periods, although the decreases noted were small. The biggest declines in population occurred in areas categorised as highly rural/remote, for the periods 1986-1991 (7 percent) and 1996-2001 (6 percent). For the periods 1981-1986 and 1991-1996, these areas witnessed small increases in population (Figure 6.5). The greatest absolute population increases were noted in main urban areas for all of the five year periods (Figure 6.6). During 1981-1986, the population of these areas increased by a total of 83,691 people. By 1991-1996 this figure had increased to 190,401, but decreased to 114,777 during the period 1996-2001. Rural areas with a high urban influence noted the next highest absolute increases in population for every time period, although the largest growth of 14,715 people during 1991-1996 was not on par with the increases noted in the main urban areas. Highly rural/remote areas witnessed the smallest absolute increases in population/the largest absolute decreases in population for every time period, with the exception of the most recent five years, where independent urban communities decreased by 7,035 people.

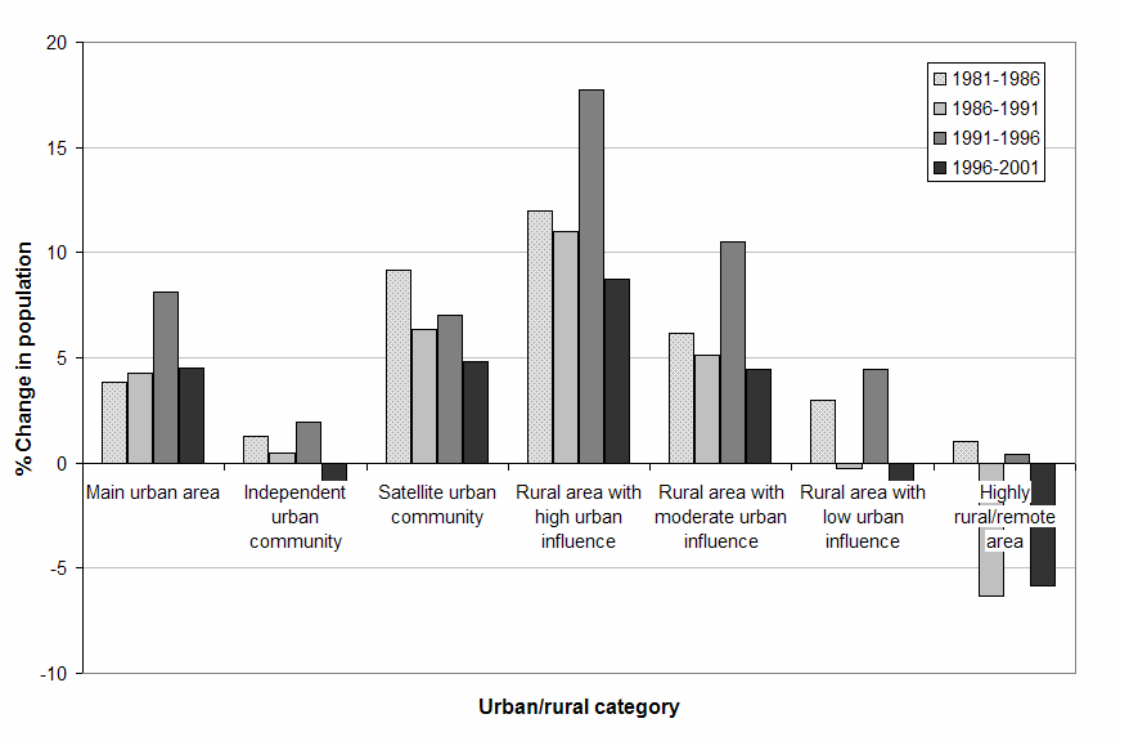

Figure 6.5: Relative population change by urban/rural category 1981-2001

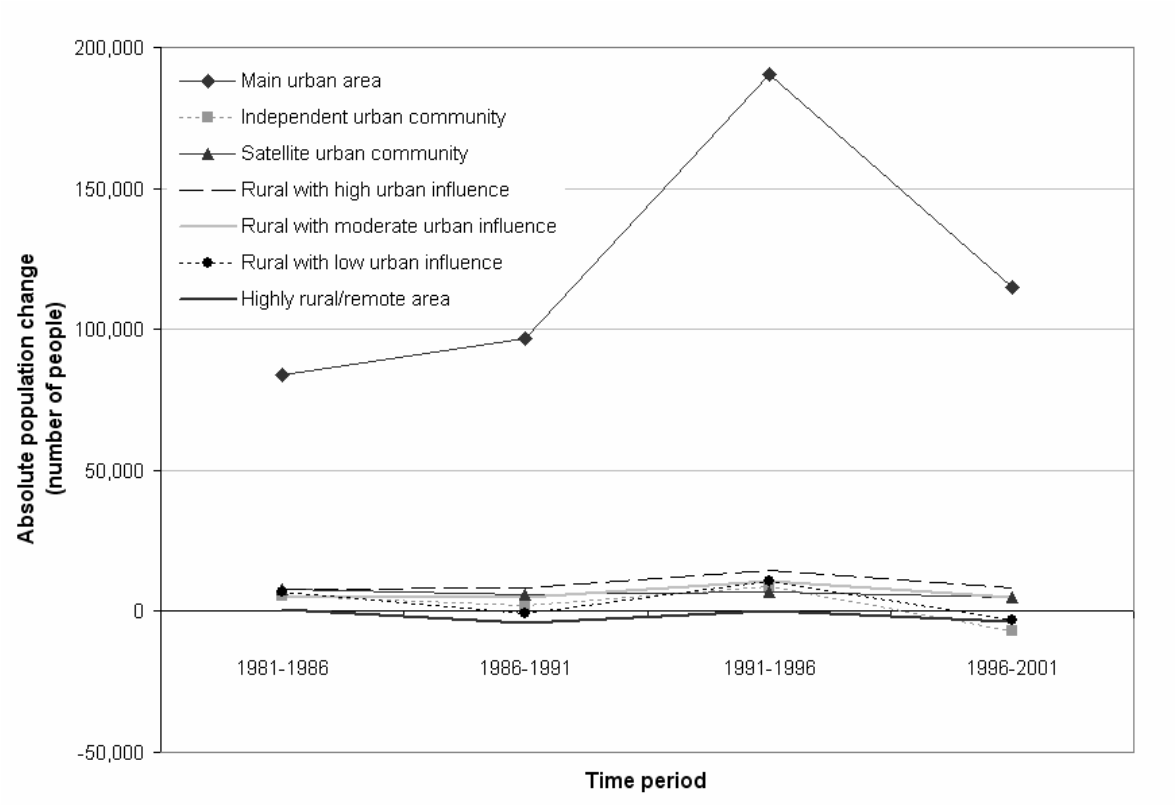

Figure 6.6: Absolute population change by urban/rural category 1981-2001

6.3 Change in the percentage of total migrants

In order to measure levels of population mixing, it is not satisfactory to simply consider net change in population over time; change in the number of migrants entering areas must also be examined. This section reports on the change in total migrants, those moving internally or from overseas since the previous census, as a percentage of the total population of each area.

6.3.1 Description of the change in the percentage of total migrants 1981-2001

At the national-level, there was an overall increase (6.92 percent) in the percentage of total migrants between 1981 and 2001 (Table 6.4). However, examining shorter time periods revealed a decrease (10.79 percent) in the percentage of total migrants in the country in the first five years of the study (1981-1986). The following five years (1986-1991) saw an increase in the percentage of migrants (9.43 percent), but this figure decreased again to just 2.13 percent by the final five years of the study period (1996-2001).

Temporal variations in this measure were also apparent at a more detailed geographical scale (CAU-level). Comparing the mean values of change in the percentage of total migrants revealed that CAUs in the first period (1981-1986) witnessed a decrease in the average percentage of migrants (-8.80%)

by area (Table 6.5). This finding contrasted strongly with the next two five year periods when mean values were both around positive seven. In fact all of the measures of central tendency for the earliest period pointed to an overall decrease in the percentage of migrants during this time. From inspection of the raw data, this result can be attributed to higher values of internal migrants in 1981 compared to 1986 in just over half of all CAUs. During this period, the largest decrease in the percentage of migrants was 71.47 percent. However, one CAU increased by 156.15 percent in this measure during this time. Large variations at the CAU-level were also noted for the decadal time periods, with the highest average increase (7.35 percent) noted between 1991 and 2001 (Table 6.5).

Table 6.4: Change in the percentage of total migrants in New Zealand between 1981 and 2001

Time Period	Change in % total migrants (%)
1981-1986	-10.79
1986-1991	9.43
1991-1996	7.23
1996-2001	2.13
1981-1991	-2.37
1991-2001	9.52
1981-2001	6.92

Table 6.5: Descriptive statistics of the change in the percentage of total migrants in New Zealand CAUs between 1981 and 2001

Time Period	Change in % total migrants (%)					
	Minimum	Maximum	Mean	Mode	Median	Standard Deviation
1981-1986	-71.47	156.15	-8.80	-44	-7.77	32.53
1986-1991	-50.50	165.04	7.77	10	7.30	14.11
1991-1996	-36.36	160.64	7.28	6	6.06	13.79
1996-2001	-89.18	74.29	0.52	2	0.58	11.86
1981-1991	-82.61	124.29	-2.27	-33	-2.95	35.14
1991-2001	-83.04	140.08	7.35	7	6.72	15.69
1981-2001	-72.83	355.82	4.47	-29	3.23	39.95

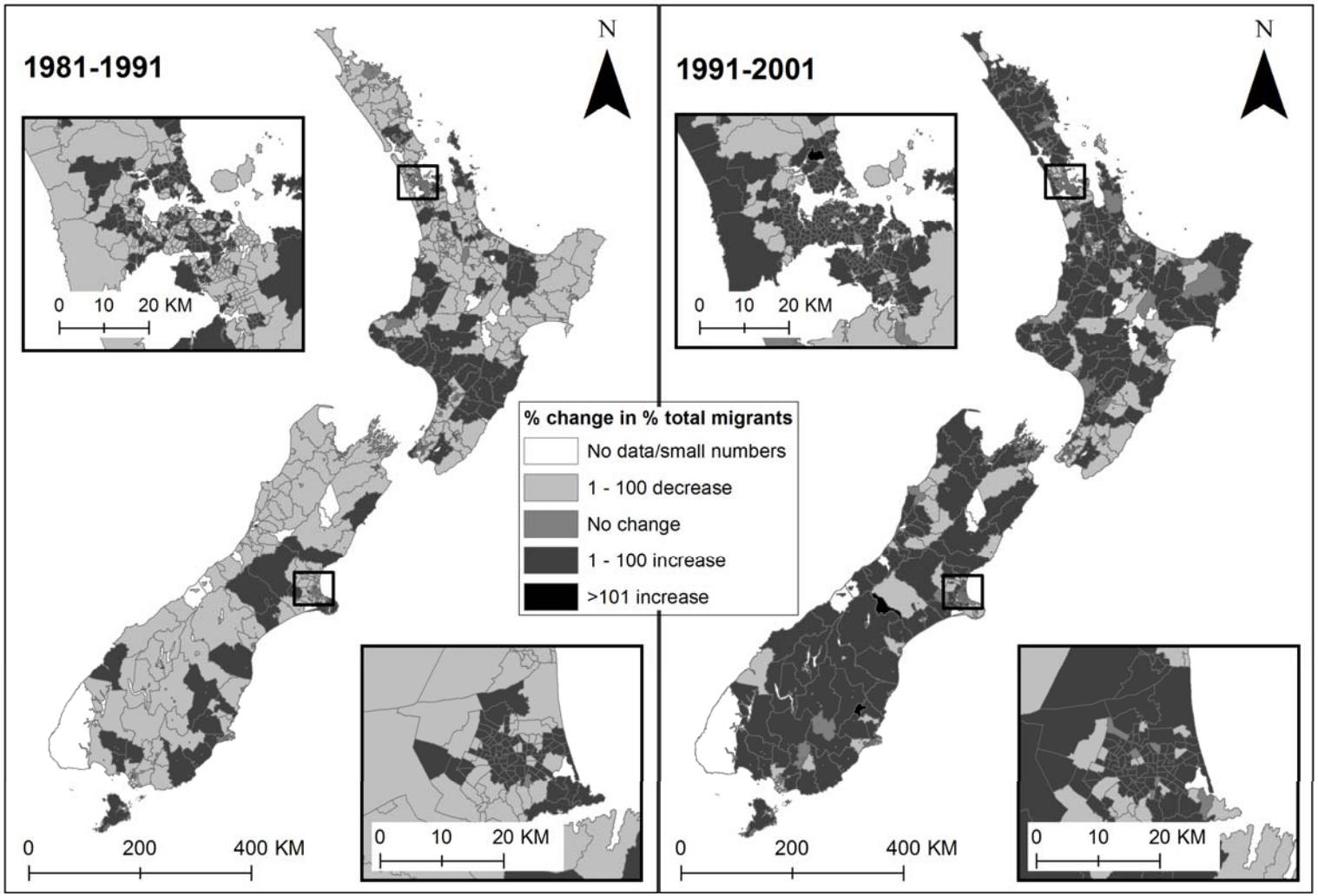

Figure 6.7: Change in the percentage of total migrants by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1981-1991 and 1991-2001

The geography of change in the percentage of total migrants at the CAU-level for the two ten year periods (1981-1991 and 1991-2001), was very distinct (Figure 6.7). During the period 1981 to 1991, an increase in the percentage of migrants occurred around the major towns and cities in the North Island especially in Auckland, Wellington, Hamilton, Rotorua, and Wanganui, and also in a belt stretching west to east, just north of the Wellington region. The majority of other areas in the North Island experienced a decrease in the percentage of total migrants over this ten year period. The majority of the South Island CAUs witnessed a decrease in the percentage of total migrants living there between 1981 and 1991. All of the CAUs along the West coast (except for the Milford Sound CAU) saw a decline in this measure, as did the majority of the northern part of the island. Growth areas included CAUs around Nelson, Blenheim, Kaikoura and its surrounding rural area, parts of central Canterbury, Christchurch and Banks Peninsula, Timaru, Oamaru, Dunedin and the far south east coast, and areas around Invercargill.

In contrast to the first ten year period, the percentage of total migrants in the majority of CAUs in the North Island increased by between one and 100 percent in the latter half of the study period (1991-2001). These increases were most common in the northern part of the island with a few exceptions, notably the rural areas surrounding the cities of Auckland, Hamilton, Tauranga and Taupo, which decreased in this measure. The majority of the South Island also witnessed an increase in the proportion of total migrants residing there between 1991 and 2001. This trend was especially evident in the southern half of the island, with exceptions in Milford Sound and rural areas around Invercargill, Dunedin and Timaru. Areas of decrease were more common in the northern half of the island but these CAUs were still interspersed with large areas of growth, for example in the far north and the Grey and Kaikoura Districts. The overall picture was dramatically different to the patterns of growth and decline depicted during the 1981-1991 period. An increase in the percentage of total migrants was noted for many more areas during the later period (1991-2001).

6.3.2 Contextual understanding of the change in the percentage of total migrants 1981-2001

When the results were disaggregated by deprivation quintile (Figure 6.8), a strong social gradient was evident for the period 1981-1986. A decrease in total migrants occurred across all of the deprivation categories for this period, with the largest decreases in the most affluent areas (more than 15 percent), and the smallest losses in the most deprived areas. Change in the percentage of total migrants in 1986-1991 and 1991-1996 did not exhibit much difference by deprivation quintile, especially in quintiles one to three. Between 1986 and 1991, marginally more growth occurred in the most deprived areas of New Zealand. During the most recent period, a positive relationship between change in the percentage of total migrants and deprivation was observed, with declines noted in the

most affluent areas and steady growth occurring in quintiles three (1.5 percent), four (4 percent) and five (7 percent). This pattern is the reverse of that shown for population change for this period.

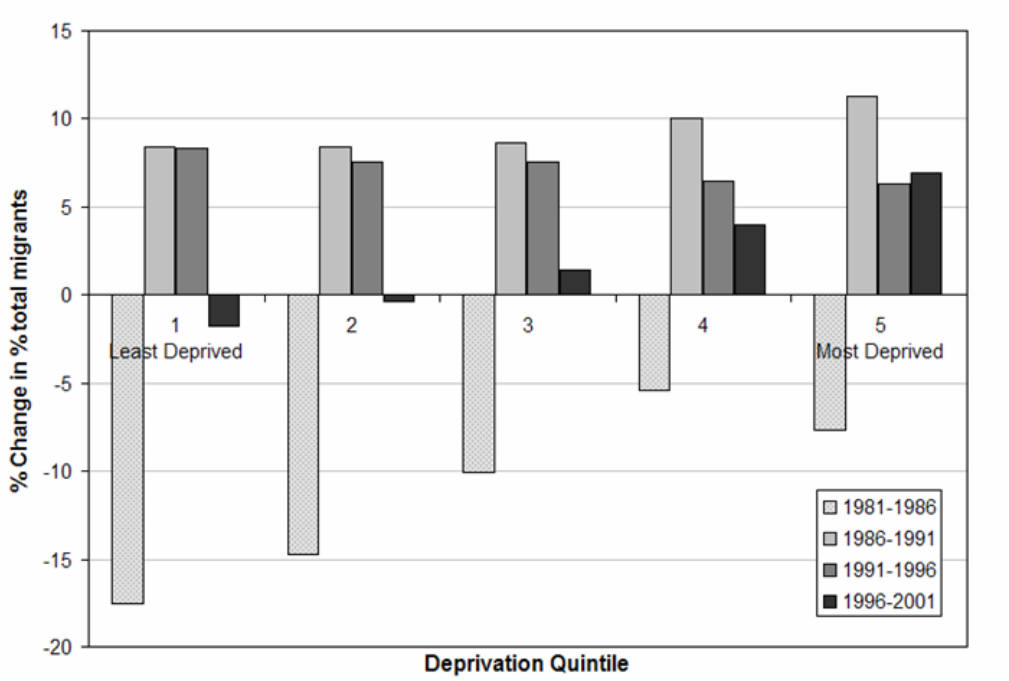

Figure 6.8: Change in the percentage of total migrants by deprivation quintile 1981-2001

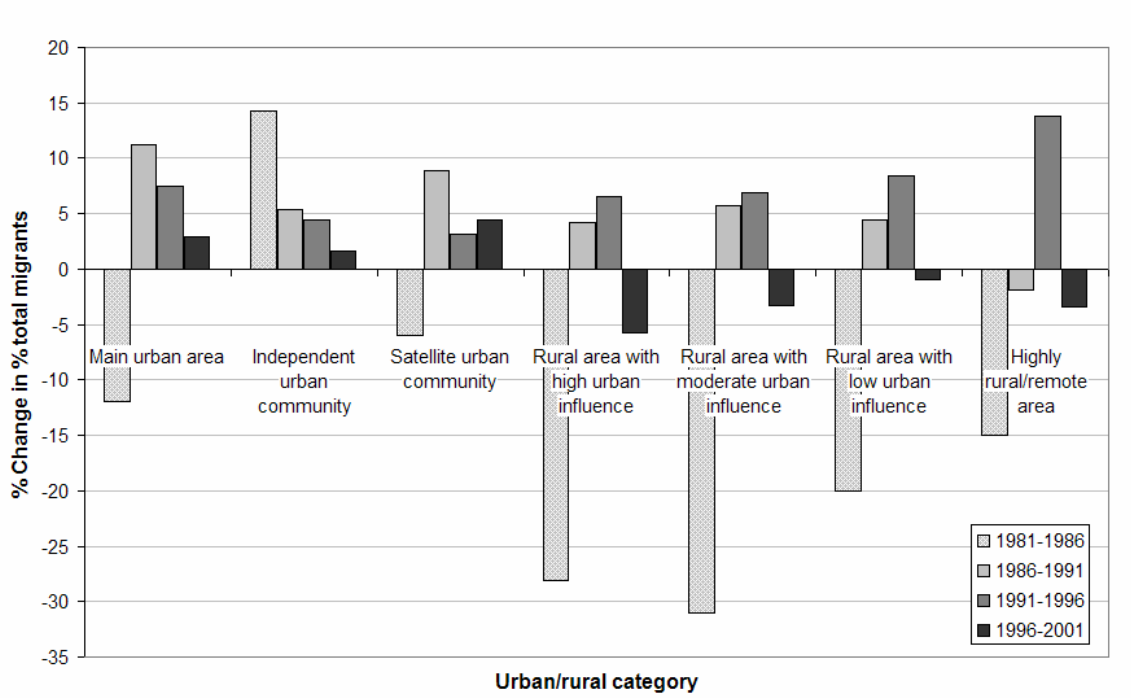

Figure 6.9: Change in the percentage of total migrants by urban/rural category 1981-2001

In terms of urban/rural trends, the only areas which experienced an overall increase in the percentage of migrants between 1981 and 1986 were areas categorised as independent urban communities (Figure 6.9). The greatest decreases seen over this period occurred in rural areas, especially those

with a moderate urban influence (over 30 percent decrease). For later periods, urban areas generally saw a rise in the percentage of total migrants (between 2 and 12 percent), as did rural areas between 1986 and 1996. Between 1996 and 2001 however, rural areas encountered a general decline in the percentage of migrants. Highly rural/remote areas experienced a decrease in the percentage of migrants during all five year periods except 1991-1996, when a 14 percent overall increase in the percentage of migrants was observed in these areas.

6.4 Change in the percentage of child migrants

As noted in Chapter 5, some researchers have postulated that infections triggering childhood ALL (Parslow et al., 2002) or type 1 diabetes (Parslow et al., 2001) are most likely to be passed between children, rather than adults. To test this theory in the New Zealand context, a measure of the change in the percentage of child migrants (internal and overseas movers aged 5 to 14 years) as a percentage of the total population, was created.

6.4.1 Description of the change in the percentage of child migrants 1981-2001

Temporal variations in the percentage of child migrants in New Zealand (Table 6.6) were similar to those noted for all-age migrants (Table 6.4). For the five year time periods for example, the earliest period (1981-1986) witnessed an overall decrease in the percentage of child migrants (22.60 percent), in contrast to the small increases noted for the later periods (3.30 to 8.91 percent). However, contrary to all-age migrants, the percentage of child migrants decreased (by 7.14 percent) at the national-level during the study period as a whole (Table 6.6).

The overall decrease in the percentage of child migrants between 1981 and 2001, and especially between 1981 and 1986, was also evident at the CAU-level (Table 6.7). For example, the average change in this measure for 1981-1986 was negative (17.12 percent), as were the other measures of central tendency. Similar to the decreases in all-age migrants, this finding can be attributed to higher numbers of internal child migrants (in 63 percent of CAUs) in 1981 when compared to 1986. Analysing the change in this measure for 1981-1991 and 1981-2001, also revealed negative mean, mode and median values, caused by elevated percentages of child migrants in 1981.

Table 6.6: Change in the percentage of child migrants in New Zealand between 1981 and 2001

Time Period	Change in % child migrants (%)
1981-1986	-22.60
1986-1991	3.30
1991-1996	8.91
1996-2001	6.63
1981-1991	-20.04
1991-2001	16.13
1981-2001	-7.14

Table 6.7: Descriptive statistics of the change in the percentage of child migrants in New Zealand CAUs between 1981 and 2001

Time Period	Change in % child migrants (%)					
	Minimum	Maximum	Mean	Mode	Median	Standard Deviation
1981-1986	-100.00	972.50	-17.12	-31	-22.52	44.13
1986-1991	-100.00	213.08	3.13	0	1.09	27.90
1991-1996	-100.00	607.87	13.66	13	8.82	39.36
1996-2001	-100.00	260.00	5.84	5	4.39	29.41
1981-1991	-100.00	347.92	-17.79	-50	-22.75	36.79
1991-2001	-100.00	634.57	16.57	14	13.89	36.83
1981-2001	-100.00	432.92	-6.74	-35	-14.17	45.56

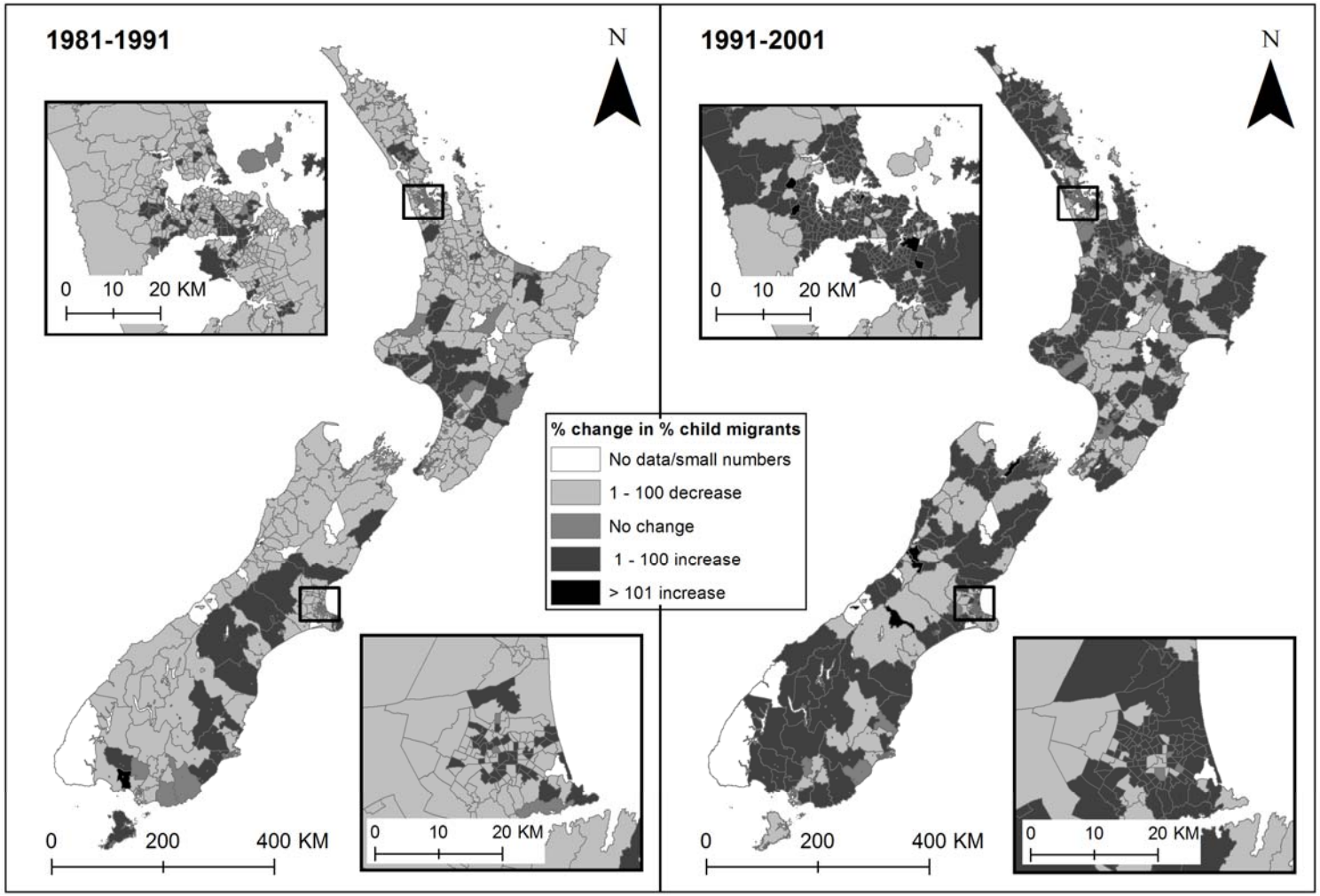

Figure 6.10: Change in the percentage of child migrants by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1981-1991 and 1991-2001

Some of the trends shown in the descriptive statistics for this measure are illustrated in Figure 6.10. The percentage of child migrants of the total population, declined over the majority of New Zealand CAUs between the years 1981 and 1991. Most CAUs in the North Island witnessed decreases in this measure, but areas of increase were noted around the Wellington area, in CAUs west, north and east of Palmerston North, south of Whakatane and around Auckland. The majority of the South Island also witnessed a decrease in the percentage of child migrants between 1981 and 1991. There was also a prominent geography to the areas that declined in this measure: the north and west coasts, as well as central areas in the south, and rural areas surrounding Christchurch all decreased in the percentage of child migrants during this period. Increases in child migrants relative to the background population were noted in central areas of the island and also parts of the south east. Increases of over 100 percent were seen in the remote rural CAUs of Fairlie in the Mackenzie TA, and Nightcaps and Fairfax in the far south of the island.

The picture of general decline in the percentage of child migrants in 1981-1991, was not continued in the following decade (Figure 6.10). Areas in the far north of the North Island generally increased in this measure between 1991 and 2001, as did CAUs in the Coromandel, and southern Auckland Regions. The eastern and western peninsulas of the North Island, and CAUs in and surrounding the Waitomo TA, also noted an increase in the percentage of child migrants between 1991 and 2001. A geographic concentration of over 100 percent growth occurred in central, western and southern Auckland. Notable increases in the percentage of child migrants between 1991 and 2001 were also seen in the South Island. Such increases were especially concentrated in the south west of the island and in CAUs to the east and south of Greymouth on the West Coast. CAUs in Christchurch, except the city centre CAUs, all increased in the percentage of child migrants during this decade. There were 33 CAUs with over 100 percent increases in this measure in the period 1991 to 2001, compared to only 10 CAUs during the earlier ten year period.

6.4.2 Contextual understanding of the change in the percentage of child migrants 1981-2001

Change in the percentage of child migrants by deprivation quintile (Figure 6.11) displayed similar trends to the change in the percentage of all-age migrants by deprivation quintile (Figure 6.8). Overall decreases were notable in all of the five deprivation quintiles for 1981-1986, with the biggest decreases in the most affluent CAUs (29 percent in quintile 1), and the smallest decreases in the most deprived CAUs (less than 20 percent in quintiles 4 and 5). As with the all-age migrants, there was little variation in this measure by deprivation deciles one to three for the 1986-1991 and 1991-1996 periods. For 1986-1991, the percentage of child migrants increased in the most deprived deciles

relative to the most affluent deciles, and the reverse was true for 1991-1996. Between 1996 and 2001, the change in the percentage of child migrants steadily increased by deprivation quintile, with the biggest rise (13 percent) occurring in the most deprived areas.

The urban/rural patterns in the change in the percentage of child migrants (Figure 6.12) were again comparable to those noted for all-age migrants (Figure 6.9). Notably, declines in all of the categories except independent urban communities were characteristic of the period 1981 to 1986, and little variation was detectable across the categories for the other time periods. Substantial increases in this measure of child migration in rural areas only occurred between 1991 and 1996.

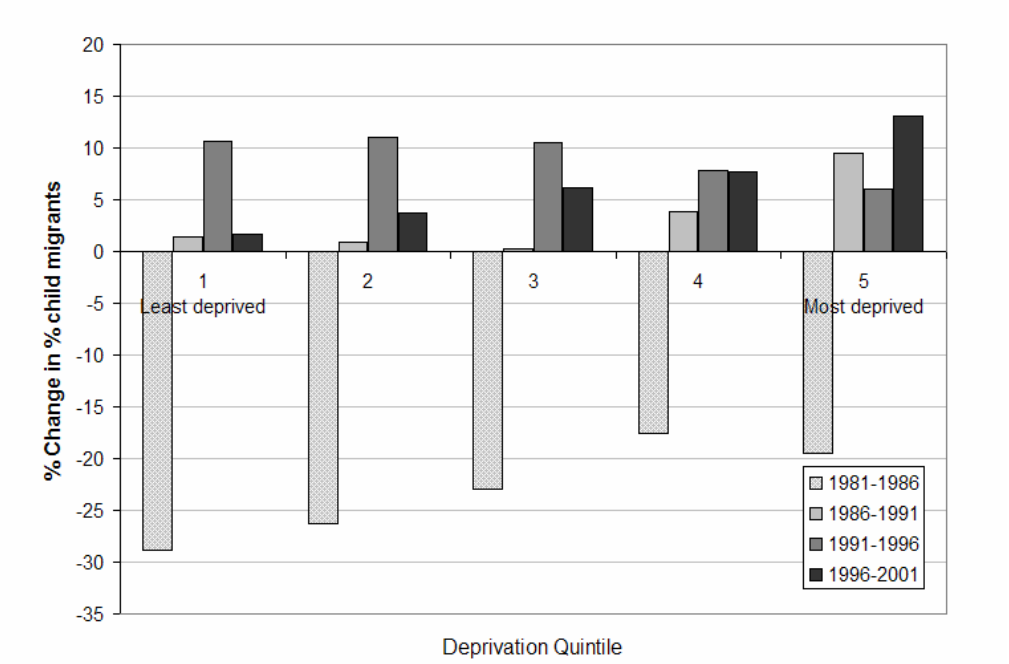

Figure 6.11: Change in the percentage of child migrants by deprivation quintile 1981-2001

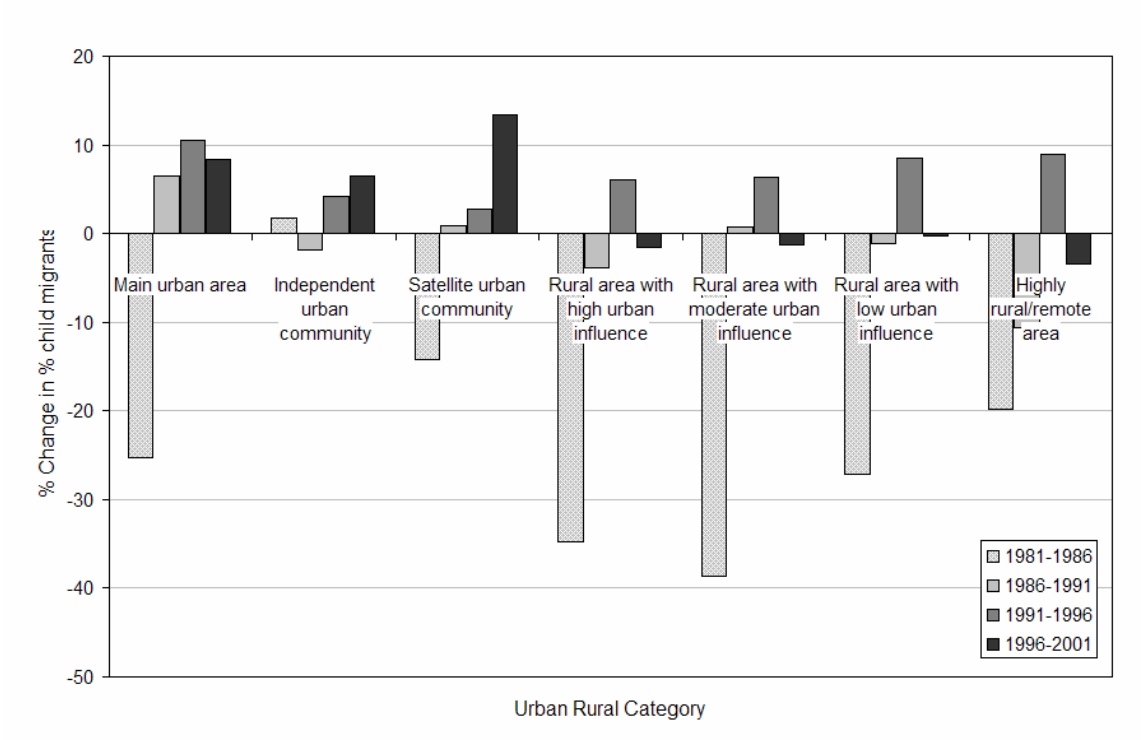

Figure 6.12: Change in the percentage of child migrants by urban/rural category 1981-2001

6.5 Change in the one year mobility percentage 1986-2001

Since the measures of change in the percentage of migrants (all-ages and child) enumerated movers in the previous five years, a measure of one year mobility was also calculated to give an indication of the number of shorter term movements. This measure was computed as the change in the percentage of people who had been in their current residence for less than one year.

6.5.1 Description of the change in the one year mobility percentage 1986-2001

Unfortunately, data for the length at usual residence was not collected for the 1981 census, so examining the change in the percentage of people who had been in their usual residence for less than one year (one year mobility percentage) before 1986 was not possible.

At the national-level, one year mobility increased by 4.22 percent between 1986 and 1991, and by a further 19.85 percent between 1991 and 1996 (Table 6.8). However, the most recent five year period (1996-2001) noted a slight decrease (1.29 percent) in this measure. Considering longer time frames revealed overall increases in the percentage of one year movers for the periods 1991-2001 (18.31 percent) and 1986-2001 (23.31 percent). At the CAU-level similar overall trends were noted (Table 6.9). For the five year periods, the greatest increase in mobility appeared to have occurred between

1991 and 1996 as this period had the highest mean (20.33 percent), mode (11 percent), median (16.90 percent), and maximum values (1,684.53 percent). As with the national-level data, slight decreases in this measure were noted between 1996 and 2001.

Table 6.8: Change in the one year mobility percentage in New Zealand between 1986 and 2001

Time Period	Change in one year mobility % (%)
1986-1991	4.22
1991-1996	19.85
1996-2001	-1.29
1991-2001	18.31
1986-2001	23.31

Table 6.9: Descriptive statistics of the change in the one year mobility percentage in New Zealand CAUs between 1986 and 2001

Time Period	Change in one year mobility percentage (%)					
	Minimum	Maximum	Mean	Mode	Median	Standard Deviation
1986-1991	-74.81	276.03	5.82	1	3.86	24.71
1991-1996	-74.20	1684.53	20.33	11	16.90	53.13
1996-2001	-75.00	211.93	-0.75	2	-1.77	21.42
1991-2001	-83.04	529.48	16.79	7	14.98	32.98
1986-2001	-72.92	474.50	21.53	13	19.14	34.66

The geography of changes in one year mobility at the CAU-level is depicted in Figure 6.13 for the five year period 1986-1991, and the ten year period 1991-2001. Between 1986 and 1991, an almost continuous area of increase in mobility occurred in the south east of the South Island, in stark contrast to areas on the West Coast to the south of Haast, which displayed a decline in mobility during these five years. In the North Island, growth areas were not as spatially contiguous but were generally noted in the northern Waikato, eastern Bay of Plenty, and southern Manawatu and Hawkes Bay regions.

In the ten year period between 1991 and 2001, there were general increases in mobility in the far south of the South Island, with notable exceptions in the Westland, Central Otago and Ashburton Districts. In the North Island, new increases in one year mobility occurred in the Wairoa District and in the majority of the Auckland main urban area. However in general, more decreases in mobility were noted in the North Island during this period compared to the previous five years (1986-1991).

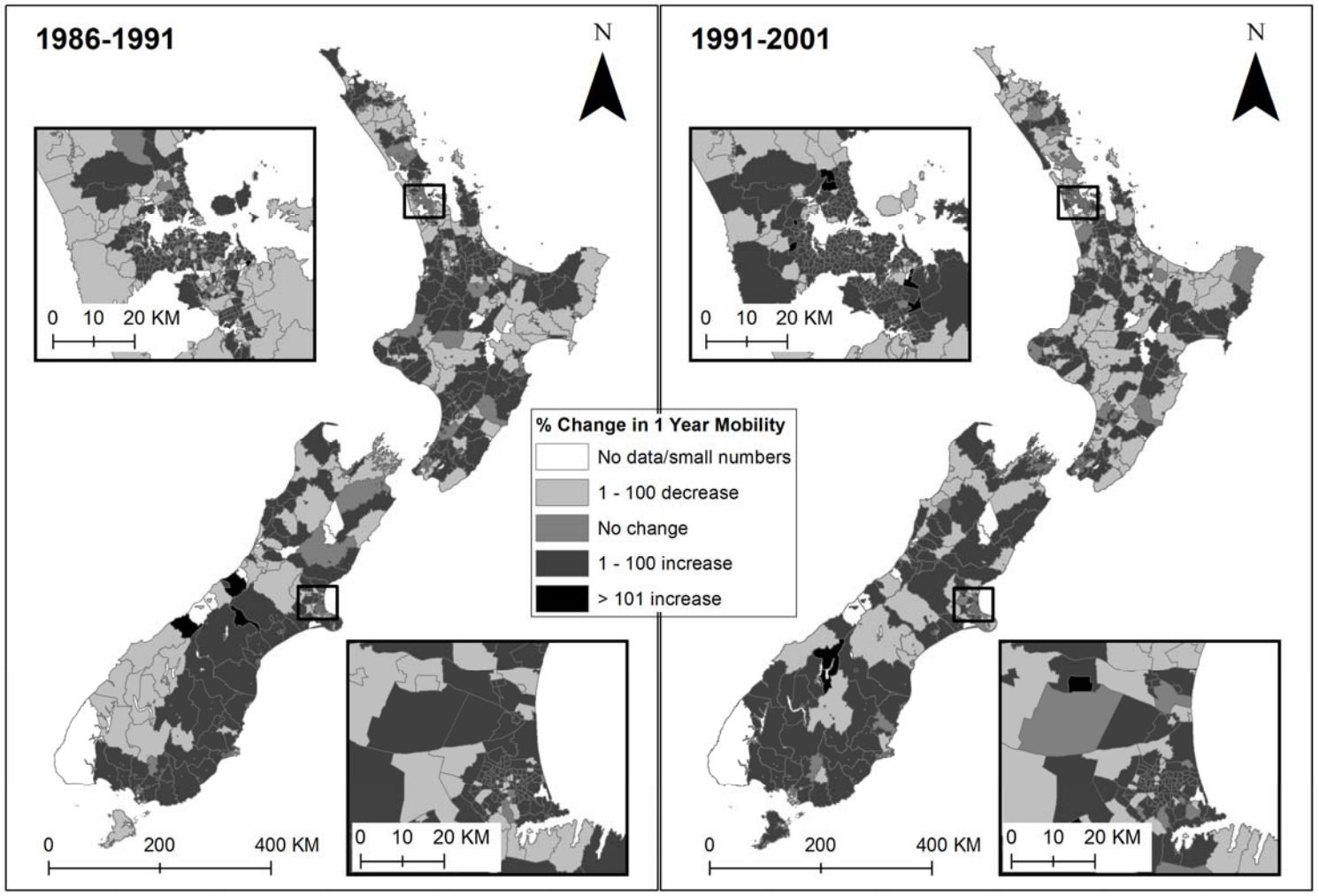

Figure 6.13: Change in the one year mobility percentage by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1986-1991 and 1991-2001

6.5.2 Contextual understanding of the change in the one year mobility percentage 1986-2001

As well as stark geographical patterns, there were also prominent trends in one year mobility by a number of contextual variables. For example, there was a clear increasing gradient between deprivation and one year mobility at the CAU-level between 1986 and 2001 (Figure 6.14). The least deprived areas of the country had the lowest percentage of one year movers compared to the most deprived fifth of areas, which had the highest percentage of one year movers. This trend was evident for every census year. Furthermore, the percentage of one year movers increased during the periods 1986-1991 and 1991-1996 in every deprivation quintile (Figure 6.15). Between 1986 and 1991, there was little difference in the increase in this measure across the five deprivation categories. However, there was a slight increasing gradient across the deprivation quintiles for the period 1991-1996. Interestingly, the percentage of one year movers decreased between 1996 and 2001 in quintiles one to four, but increased in quintile five (most deprived).

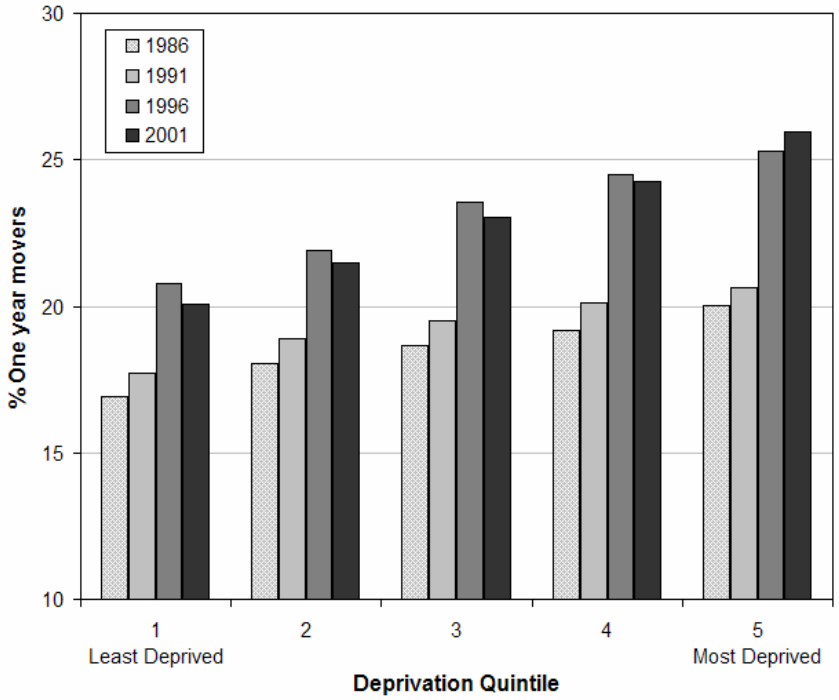

Figure 6.14: One year mobility percentages by deprivation quintile 1986-2001

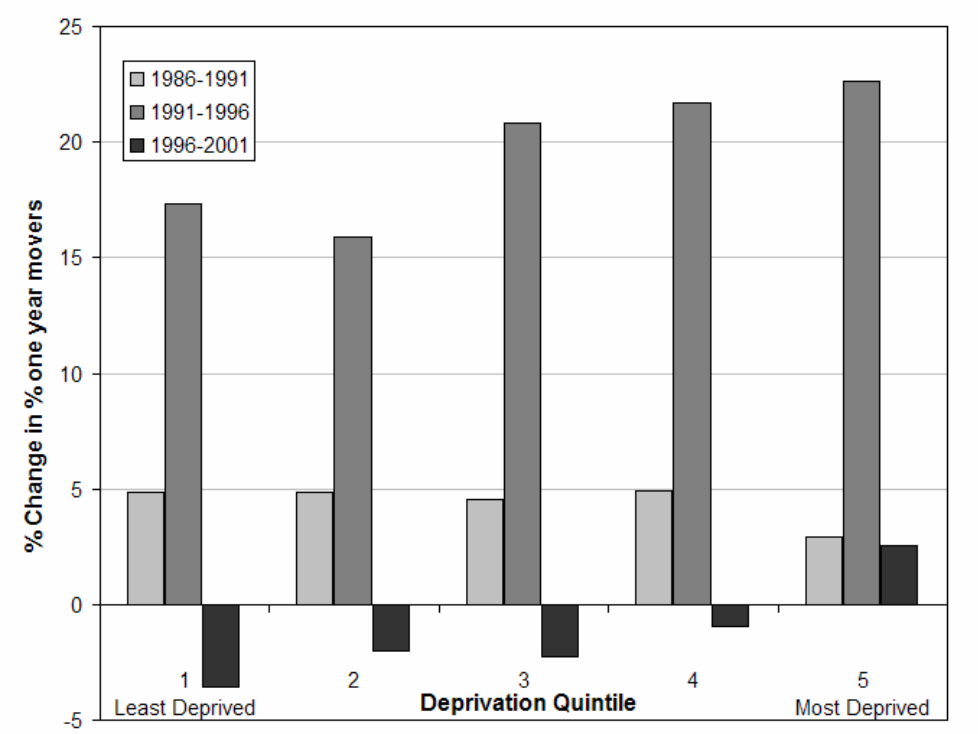

Figure 6.15: Change in the one year mobility percentage by deprivation quintile 1986-2001

The percentage of one year movers was generally higher in urban compared to rural areas of New Zealand, and this trend was especially marked for the most recent years (1996 and 2001) (Figure 6.16). Considering change over time, in the earliest period (1986-1991) there was an increase in the percentage of one year movers in the most urban and most rural areas and small decreases/no change in the other categories (Figure 6.17). The one year mobility percentage increased in every urban/rural category for the period 1991-1996, but these increases were largest in the urban areas of the country, especially satellite urban communities (27 percent). However, a 14 percent increase in mobility was also seen in highly rural/remote areas during this period. Between 1996 and 2001, there was a decrease in mobility in all of the urban/rural categories, with the biggest decline noted in rural areas with a high urban influence.

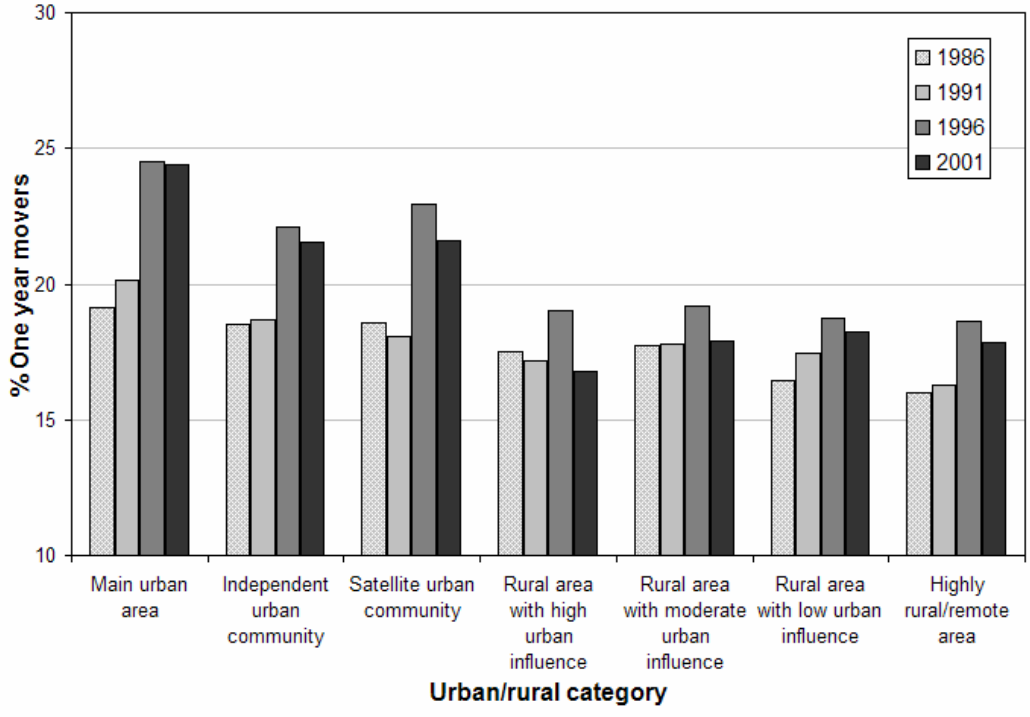

Figure 6.16: One year mobility percentages by urban/rural category 1986-2001

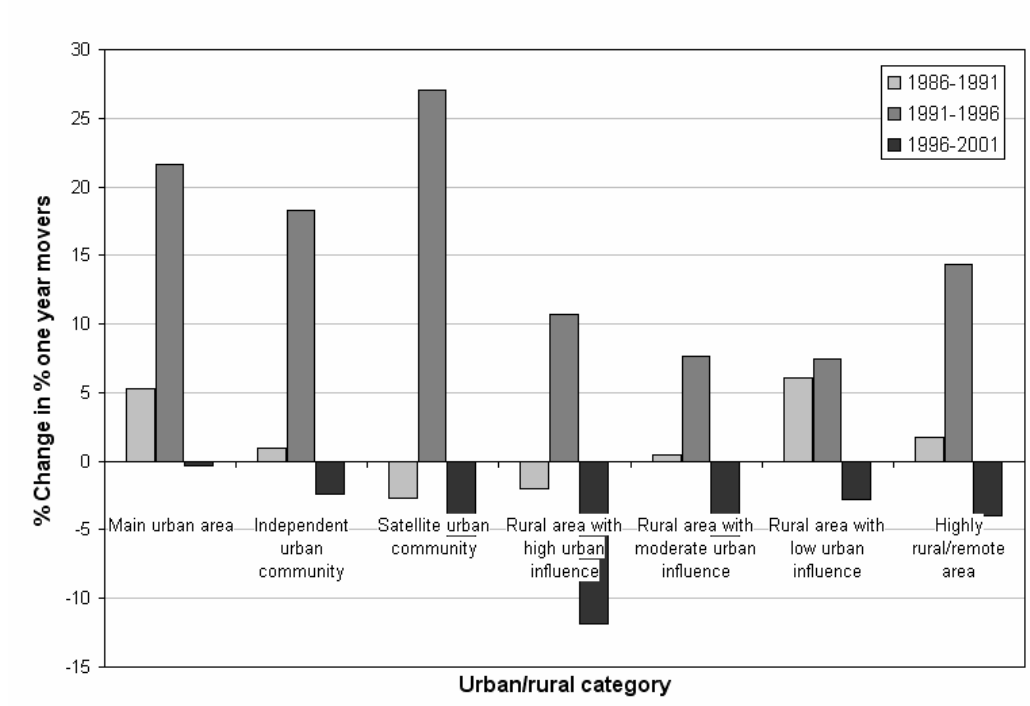

Figure 6.17: Change in the one year mobility percentage by urban/rural category 1986-2001

6.6 Change in migrant diversity

All of the population mixing measures considered thus far have been proxy measures for the *number* of infections introduced to an area. The next measure to be examined concerns the diversity of

origins of the various migrants entering an area, and hence acts as a measure for the potential *range* of infections introduced by these migrants. At the CAU-level, migrant diversity was calculated using the Shannon index of diversity and measured the extent to which the total incomers to an area (from elsewhere in New Zealand and overseas) were distributed among their origin areas (the TA within New Zealand or overseas country). At the national-level, migrant diversity was assessed by considering the origins of overseas migrants only.

6.6.1 Description of the change in migrant diversity 1981-2001

For New Zealand as a whole, the diversity of overseas migrants increased by 25.22 percent between 1981 and 2001 (Table 6.10). However, this figure hides shorter term variation. For example, between 1981 and 1986 there was a slight decrease in this measure (0.15 percent), yet over the next five years (1986-1991), overseas migrant diversity increased by 14.23 percent. Small increases were also noted for the periods 1991-1996 (4.19 percent) and 1996-2001 (5.38 percent). The ten year period with the greatest increase in this measure was 1981-1991 (14.06 percent).

Table 6.10: Change in the diversity of overseas migrants in New Zealand between 1981 and 2001

Time Period	Change in overseas migrant diversity (%)
1981-1986	-0.15
1986-1991	14.23
1991-1996	4.19
1996-2001	5.38
1981-1991	14.06
1991-2001	9.79
1981-2001	25.22

The average change in migrant diversity (internal and overseas) at the CAU-level was relatively small for all of the time periods (Table 6.11). However, the minimum, maximum and standard deviation values indicated that there was a great deal of variation between CAUs. The greatest variation and highest average growth in migrant diversity occurred between 1981 and 1986, with a mean growth of 3.57 percent, a maximum growth of over 600 percent, and a standard deviation of almost 30. For the ten year periods, the average change in diversity was marginally higher for the first ten years (3.74 percent), compared to the second ten years (3.51 percent). However, the first decade (1981-1991) noted a greater variation in migrant diversity change values.

Table 6.11: Descriptive statistics of the change in migrant diversity in New Zealand CAUs between 1981 and 2001

Time Period	Change in migrant diversity (%)					
	Minimum	Maximum	Mean	Mode	Median	Standard Deviation
1981-1986	-100.00	683.34	3.57	0	-0.41	29.91
1986-1991	-67.12	275.43	0.50	0	0.30	14.42
1991-1996	-54.08	164.64	2.61	0	1.05	13.23
1996-2001	-72.98	210.82	1.53	1	0.91	14.95
1981-1991	-100.00	715.62	3.74	1	-0.06	31.69
1991-2001	-60.95	346.04	3.51	2	1.76	18.61
1981-2001	-60.45	1811.38	7.69	4	1.66	60.63

A large number of CAUs declined in migrant diversity between 1981 and 1991 (Figure 6.18). The majority of CAUs in the North Island of New Zealand decreased (or noted little change) in migrant diversity during this time. However these areas are interspersed with CAUs whose diversity grew, resulting in a somewhat random distribution of high and low values across the North Island. Increases of over 100 percent occurred in the outskirts of Auckland, Hamilton, Tauranga, and Whakatane. The South Island witnessed an overall decline in migrant diversity on the West Coast from just south of Hokitika to as far south as Fjordland. Much of inland Canterbury also recorded a decrease in the diversity of incoming migrants over this period. There was a general pattern of increased diversity in the south east section of the island (with a few exceptions), whereas the northern part of the island displayed much more spatial heterogeneity. Areas which increased in diversity by over 100 percent included CAUs on the outskirts of Nelson, Christchurch, and Dunedin, and in areas around Invercargill.

As with the earlier period, most areas in the North Island of New Zealand either decreased or experienced little change in migrant diversity scores between 1991 and 2001 (Figure 6.18). However, there were some notable anomalies to this overall trend, including growth in the majority of the Far North, Opotiki and Gisborne TA areas, and also CAUs which fall within the mid-central district health board area. Increases in diversity of over 100 percent occurred in CAUs on the outskirts of Auckland, Hamilton and Gisborne. The majority of the southern half of the South Island experienced either an increase or little change in the diversity of incoming migrants between 1991 and 2001. Exceptions to this general pattern included the Haast CAU, areas in Central Otago and areas around both Invercargill and Dunedin, whose diversity reduced by between negative two and negative 15 percent. The northern half of the South Island had a relatively random spatial distribution of change in diversity scores over this period. Overall, it can be noted that the majority of areas increased in diversity between the 1980s and 1990s, particularly in the South Island of New Zealand.

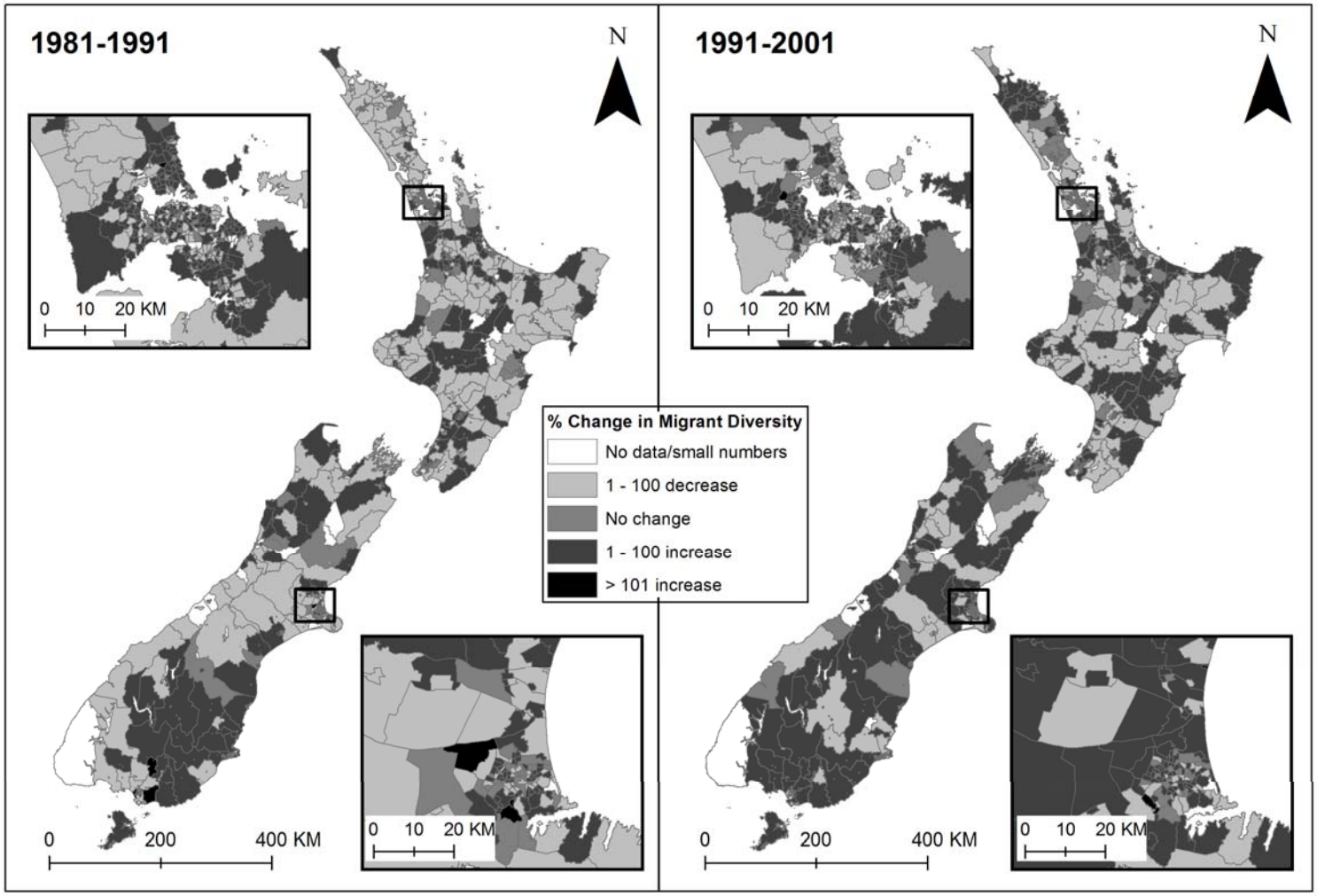

Figure 6.18: Change in migrant diversity by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1981-1991 and 1991-2001

6.6.2 Contextual understanding of the change in migrant diversity 1981-2001

In order to determine what types of area were receiving migrants from a wide range of origins compared to those receiving a less diverse pool of migrants, average diversity scores were summarised by both deprivation quintile and urban/rural status. Average migrant diversity scores were found to be consistently higher in the most deprived areas when compared to the most affluent areas of New Zealand (Figure 6.19). However, this gap narrowed over time. The ratio between the scores in the best and worst off areas decreased from 1.18 in 1981, to 1.08 by 2001. The average diversity of migrants in deprivation quintile one areas increased during every time period, with the largest increase of approximately four percent noted between 1981 and 1986 (Figure 6.20). The most deprived areas of New Zealand generally witnessed a decrease in average migrant diversity, with the exception of the period 1991-1996.

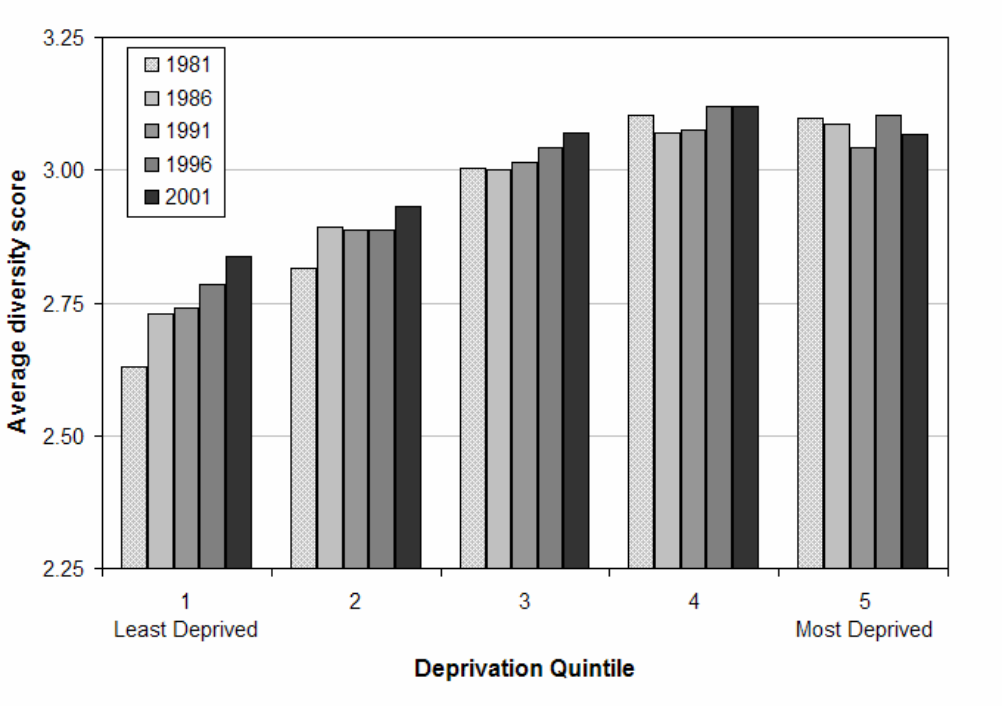

Figure 6.19: Average migrant diversity score by deprivation quintile 1981-2001

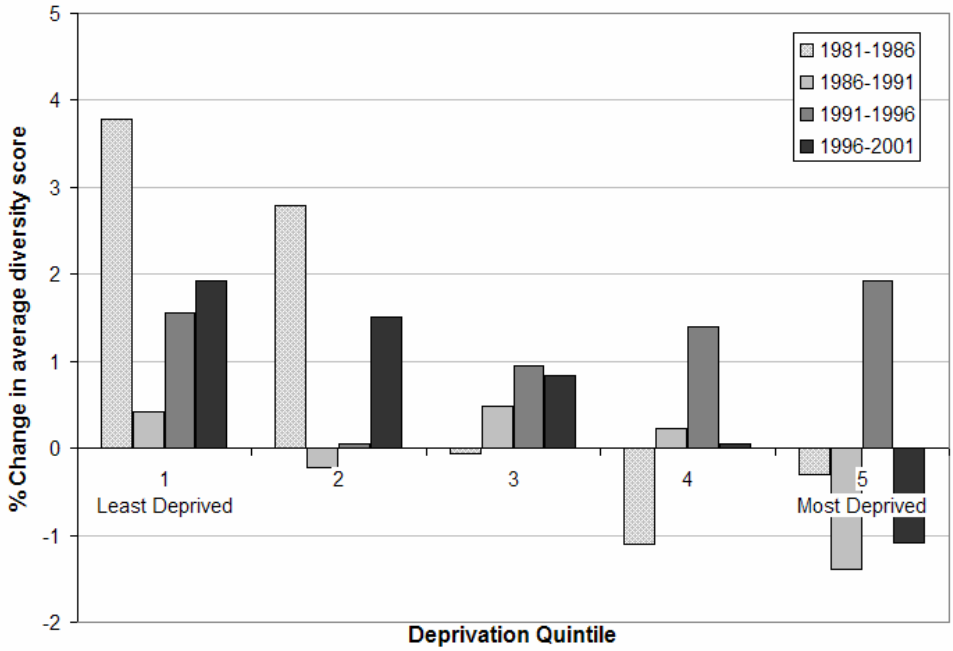

Figure 6.20: Change in average migrant diversity scores
by deprivation quintile 1981-2001

Average migrant diversity was also consistently higher in the urban areas of New Zealand for all years (Figure 6.21). The main urban areas and independent urban communities had the highest average diversity scores of over three for every year. The lowest migrant diversity scores (less than 2.50) were noted in rural areas with a high urban influence. Change in this measure by five year periods is depicted in Figure 6.22 which reveals few obvious trends. The earliest period generally saw little change or a decline in diversity in urban areas, and increases in rural areas, especially those with a high urban influence. Between 1986 and 1991, migrant diversity decreased in every category except main urban areas. For the later two periods (1991-1996 and 1996-2001), there was little change in the average migrant diversity score across the urban/rural spectrum.

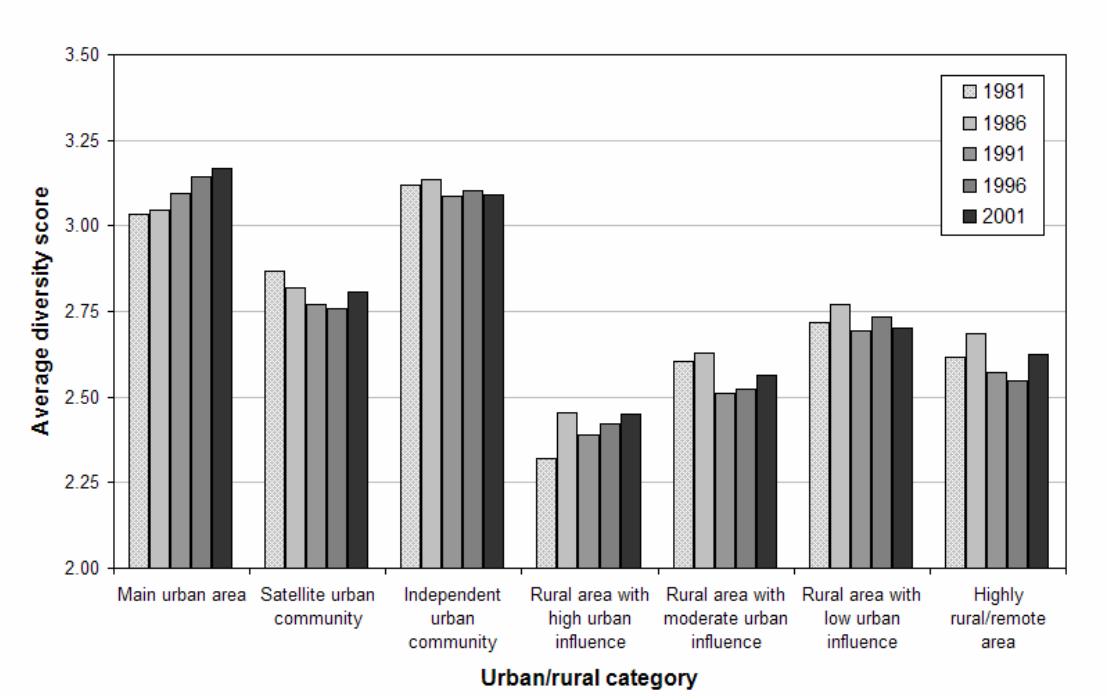

Figure 6.21: Average migrant diversity scores by urban/rural category 1981-2001

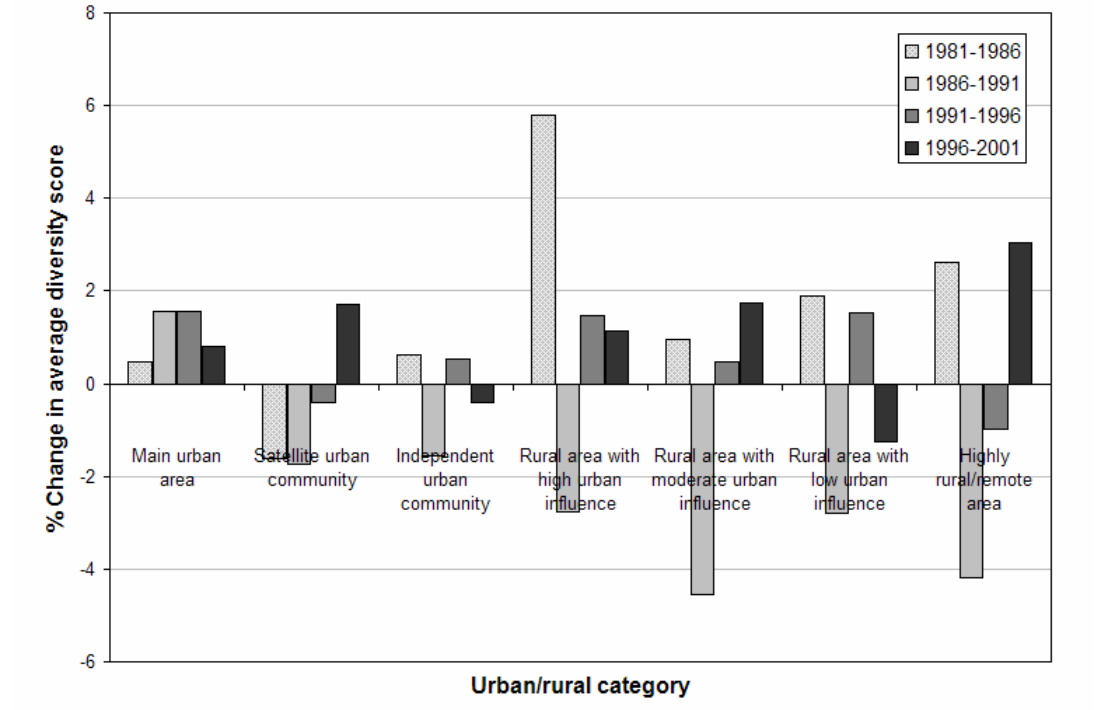

Figure 6.22: Change in average migrant diversity scores by urban/rural category 1981-2001

6.7 Population mixing categories

While the volume of in-migration acts as an alternative measure for the possible number of infections entering an area, and the diversity of origins of in-migrants can be utilised to represent the potential

range of infections introduced; it is also possible that these measures in combination will have implications for the spread of infections (Cox, 2007). As a result, an additional categorical variable was created following Cox (2007) to investigate these likely interactions. Each CAU in New Zealand was thus assigned to one of four categories: low in-migration and low migrant diversity; low in-migration and high migrant diversity; high in-migration and low migrant diversity; and high in-migration and high migrant diversity.

6.7.1 Description of the population mixing categories 1981-2001

In 1981, the majority of CAUs in New Zealand (39.35 percent) were classed as having low migration and high migrant diversity (Table 6.12). In contrast, the lowest population mixing category of low migration and low migrant diversity was the most frequently occurring category for all of the later years.

Table 6.12: Population mixing category by the percentage of total CAUs and year 1981-2001

Population Mixing Category	Category Abbreviation	% of total CAUs				
		1981	1986	1991	1996	2001
1. Low in-migration & low diversity	LMLD	13.61	30.72	34.38	35.13	36.03
2. Low in-migration & high diversity	LMHD	39.35	30.60	23.11	24.59	20.21
3. High in-migration & low diversity	HMLD	36.39	13.25	15.15	13.52	13.65
4. High in-migration & high diversity	HMHD	10.65	25.43	27.36	26.76	30.11

The greatest changes between categories were observed for the period 1981-1986 (Figure 6.23). For example during this time, the number of CAUs in the highest population mixing category (high migration and high diversity) increased by nearly 140 percent, whilst those in the lowest population mixing category also increased by 125 percent. The number of CAUs in each of the middle two population mixing categories decreased during this time. Changes for the other five year periods were not as substantial as those shown by the first time-span, but were nonetheless notable. During 1986-1991 there were increases in the number of CAUs in all of the categories, except for low migration and high diversity, which decreased by 24 percent over this period. Between 1991 and 1996, decreases were experienced in the high population mixing categories (3 and 4), whilst increases were only noted in the low population mixing categories (1 and 2). The most recent period showed a 13 percent increase in the number of CAUs in the highest population mixing category (4).

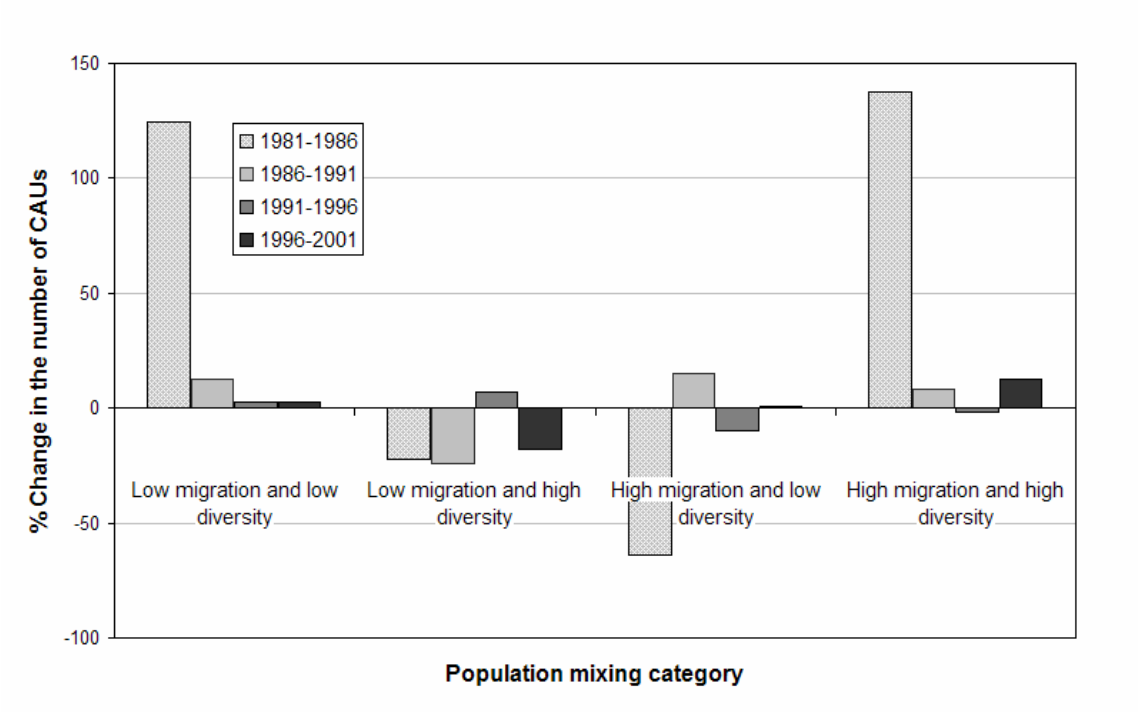

Figure 6.23: Percentage change in the relative importance of the population mixing categories by five year periods, 1981-2001

6.7.2 Geography of the population mixing categories 1981-2001

Significant patterns were detectable when comparing the spatial distribution of the various population mixing categories between 1981 and 1986 (Figure 6.24). As observed by comparing the modal classes over time (Table 6.12), there were many more CAUs categorised as having high in-migration and low diversity (HMLD) in 1981. These CAUs were especially concentrated in the north of the North Island and on the West Coast of the South Island. All of the main city centres in 1981 on both islands were classed as having both high in-migration and high diversity (HMHD). There also appeared to be a cluster of this high population mixing category in the central parts of the North Island, especially around Hamilton and Lake Taupo. In the South Island, Milford, Glenorchy, Queenstown and Moteuka were all HMHD areas. There were very few areas in the North Island which experienced low population mixing during this time (LMLD). More LMLD areas were present in the South Island, and were randomly distributed on the eastern half of the island. By 1986, the population mixing categories appeared to be more spatially contiguous, and there were significant decreases in the HMLD category CAUs. The majority of the North Island was categorised as having low in-migration and high diversity (LMHD), and increased in HMHD in the centre and outskirts of Auckland. The pattern of high population mixing in the centres of all of the major cities and around Hamilton and Lake Taupo remained. An increase in low population mixing (LMLD) was seen in the West Cape and southern Gisborne TA of the North Island, and in the far south of the South Island. In

1986 there was a larger continuous area of HMHD in the south west of the South Island; from as far south as Lumsden up the West Coast to Haast, and also including Queenstown, Wanaka and the two glacier settlements of Franz Joseph and Fox Glaciers. Increases in this high level of population mixing were also noted in Nelson and Blenheim.

By 1991 there was a further substantial increase in the number of low population mixing (LMLD) CAUs across both islands (Figure 6.25). In addition, many CAUs that were classed as LMLD or LMHD in 1986 seemed to have remained in these categories by 1991. The number of CAUs classed as having high population mixing (HMHD) also appeared to be fewer in 1991 compared to 1986. HMHD areas that remained by this time included all of the major city centres on both islands, CAUs around Lake Taupo in the centre of the North Island, and the Milford and Glenorchy CAUs in the South Island.

Comparing the 1991 trends to those in 1996 (Figure 6.26), there was a growth in CAUs categorised as having high population mixing (HMHD) in the Rodney District and central areas of the North Island around Lake Taupo. In the South Island, increases in this category occurred in the north of the island, in the Queenstown-Lakes TA, and also on Stewart Island off the far south coast. CAUs in the main urban areas were also classed as HMHD, as in previous years. Other noticeable patterns between 1991 and 1996 included an increase in CAUs experiencing low population mixing (LMLD) in the West Cape and southern tip of the North Island, and a general increase in population mixing category on the West Coast of the South Island where LMLD was present in 1991.

The spatial distribution of population mixing categories in 2001 was similar to that for 1996 (Figure 6.27), although some differences were visible. For example, by 2001 there were some further increases in the number of CAUs classed as LMLD. Such areas were situated in the Far North and Whangarei, on the outskirts of the main urban areas of the Auckland Region, and in the Hawkes Bay area of the North Island, and also in parts of the West Coast of the South Island. In addition, some areas of high population mixing (HMHD) in the centre of the North Island, the top of the South Island and in the Queenstown-Lakes TA in the South Island, had decreased in this measure since 1996. Important similarities for both years included the general concentration of high population mixing areas in the centres of the main cities, in the central North Island, in the Queenstown-Lakes TA, and in Stewart Island.

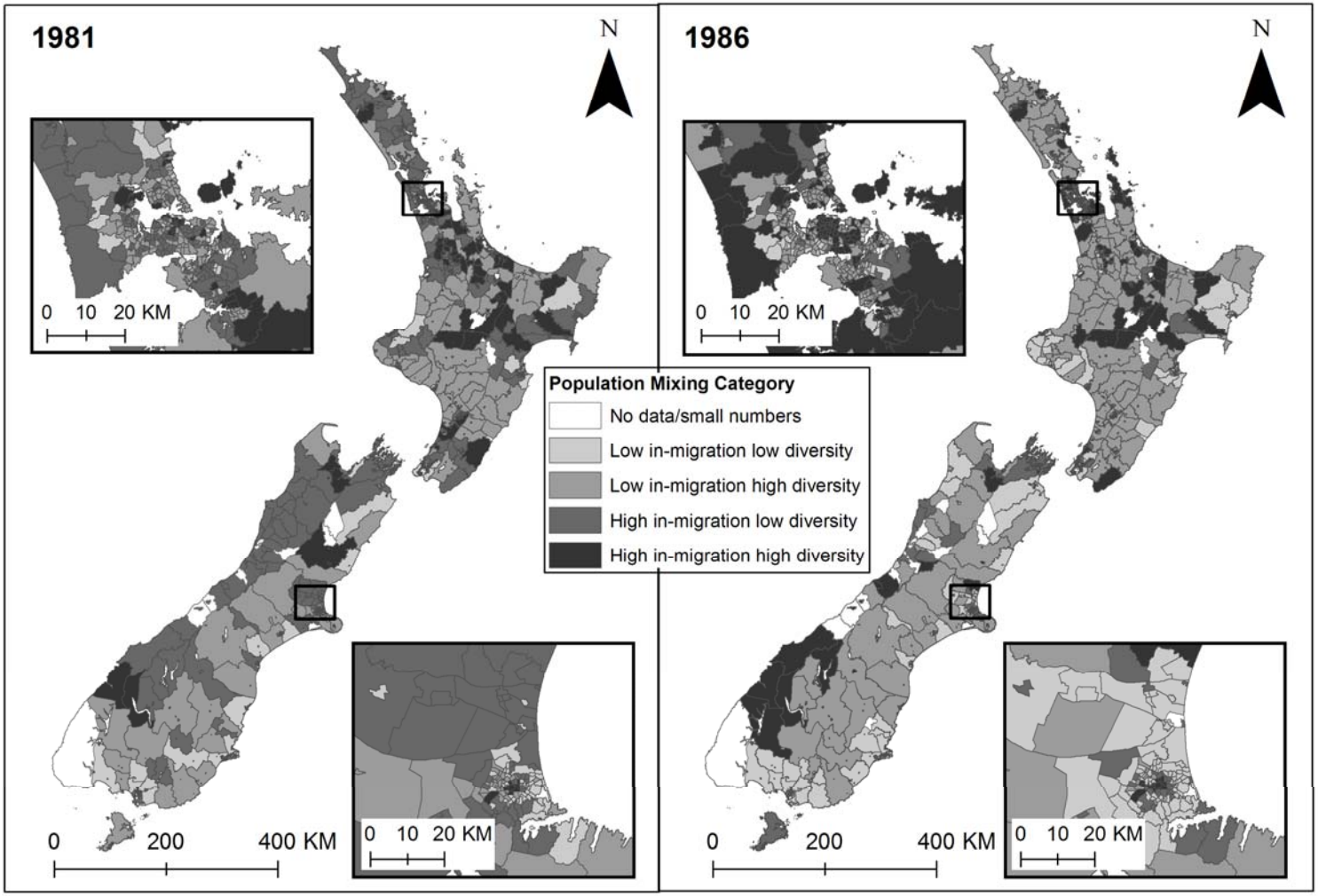

Figure 6.24: Population mixing category by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1981 and 1986

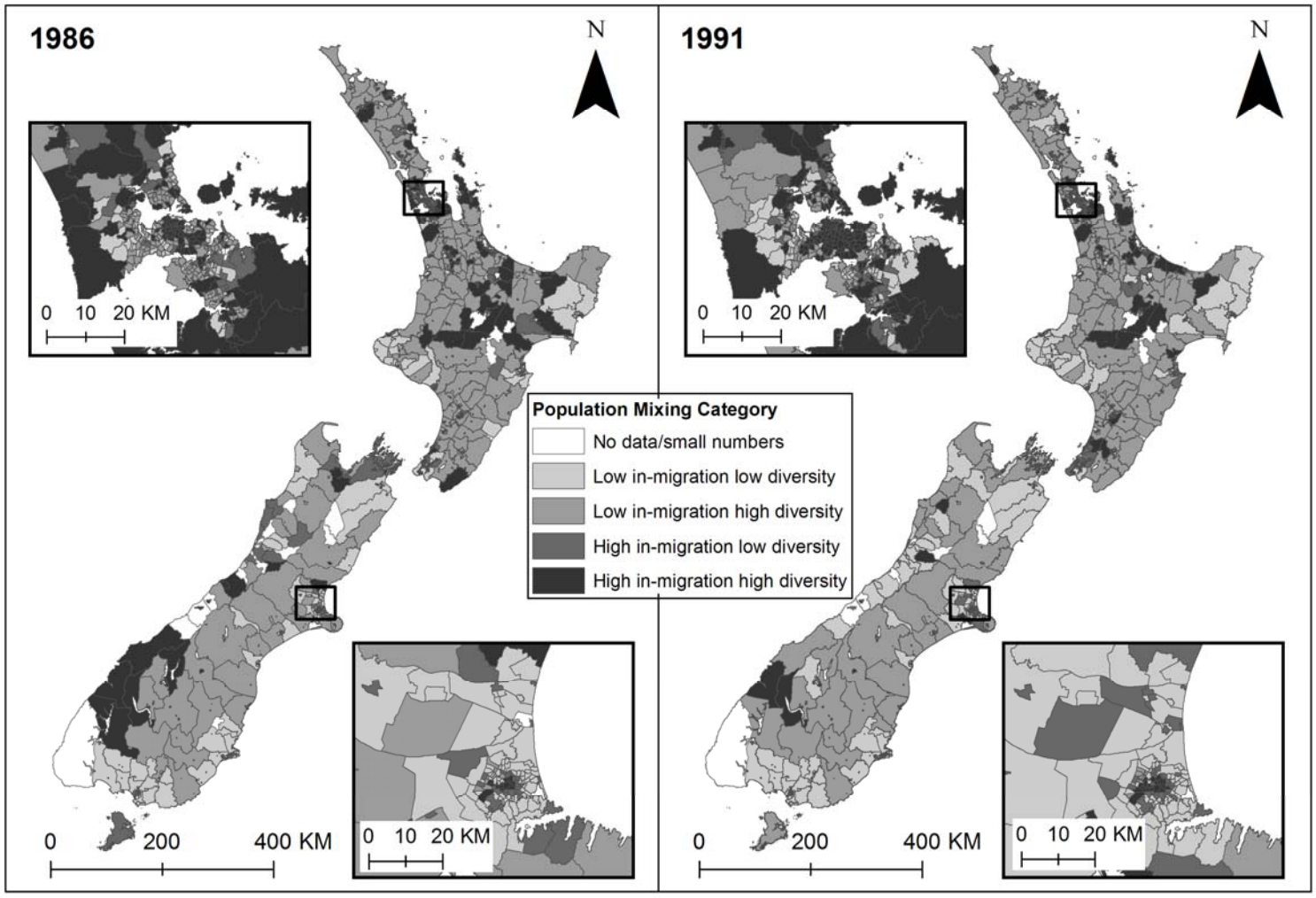

Figure 6.25: Population mixing category by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1986 and 1991

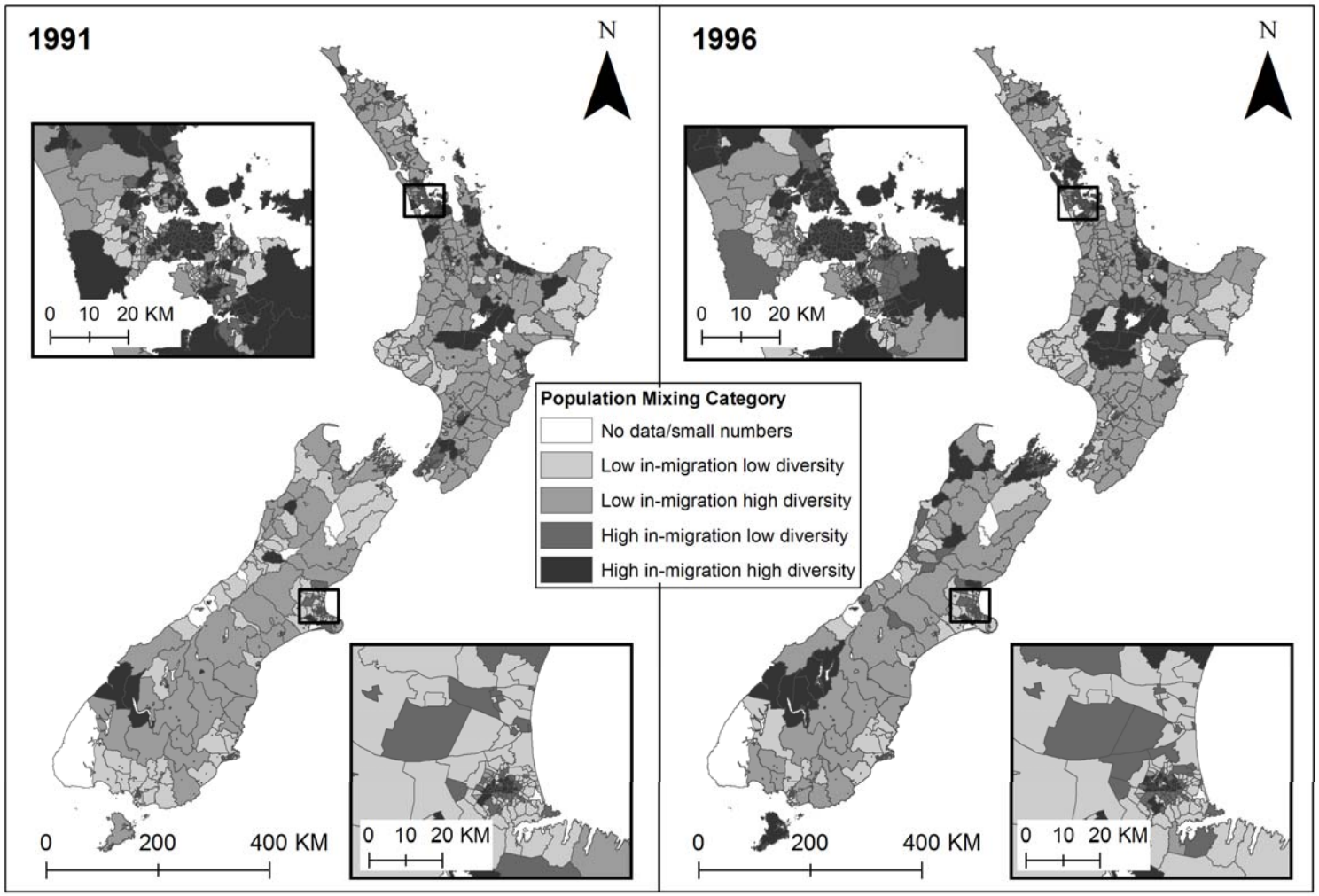

Figure 6.26: Population mixing category by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1991 and 1996

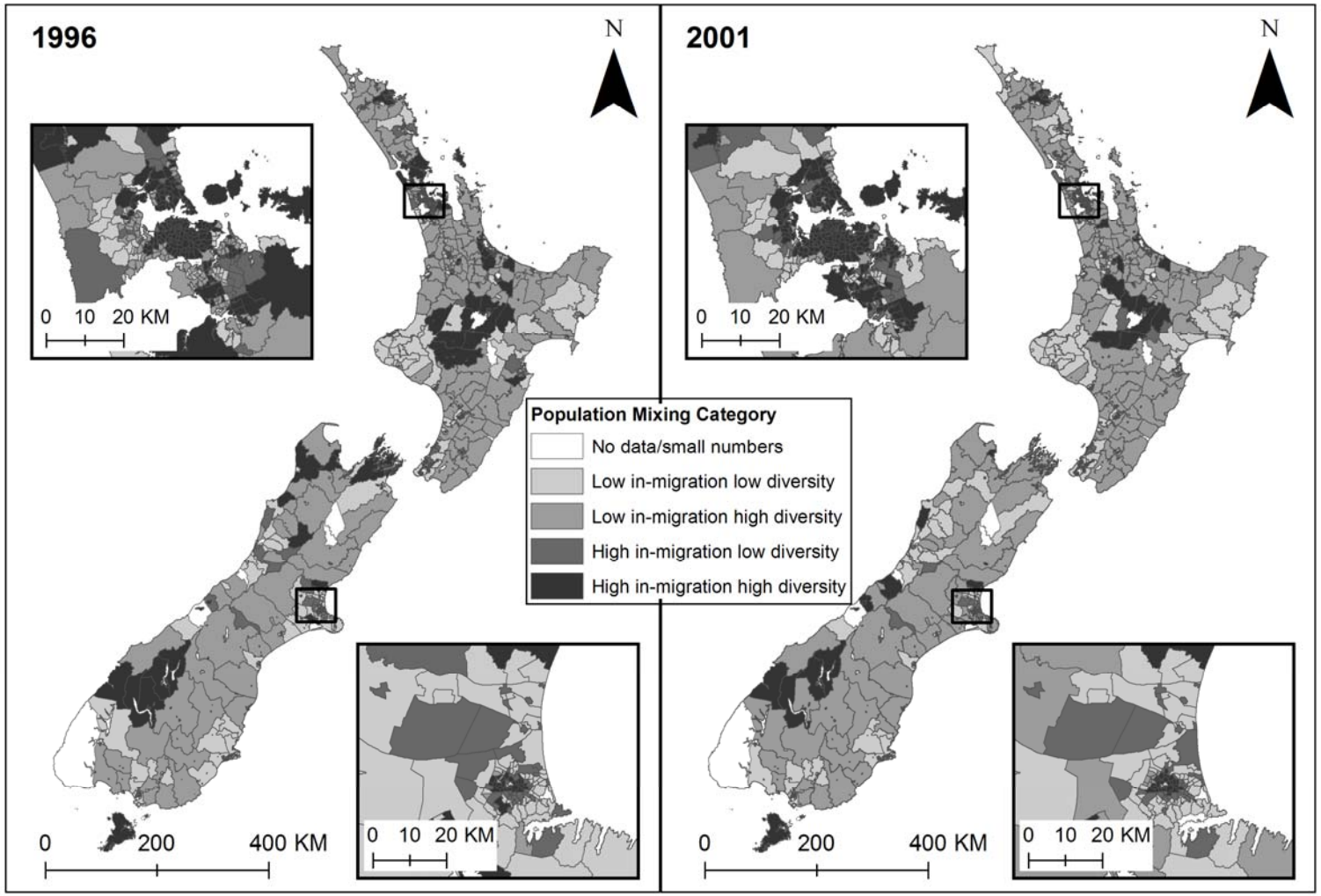

Figure 6.27: Population mixing category by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1996 and 2001

6.7.3 Contextual understanding of the population mixing categories 1981-2001

As well as a distinct geography, a number of contextual patterns were noted with the various population mixing categories at the CAU-level. For example, at the beginning of the study period (1981), CAUs classed as having the lowest levels of population mixing (LMLD and LMHD) had the highest mean population counts and mean population densities (Table 6.13). However by the middle (1991) and end of the study period (2001), the highest population mixing category (HMHD), had the highest mean population count and density compared to the other population mixing categories (Tables 6.13-6.15). Trends by deprivation quintile were ascertained by examining the modal deprivation quintile for each population mixing category. In 1981, there was no real pattern of increasing/decreasing average deprivation by population mixing category. HMLD areas tended to be the most affluent areas (quintile 1), whereas the other three population mixing categories (LMLD, LMHD and HMHD) were more likely to be deprived areas (quintiles 4 and 5) (Table 6.13). However, by 1991 the lowest population mixing areas (LMLD) tended to be the most affluent areas (quintile 1). The modal deprivation quintile for the LMHD population mixing category was quintile 2 (Table 6.14). This pattern remained in 2001 (Table 6.15). Moreover by 2001, the highest population mixing categories (HMLD and HMHD) were more likely to be deprived areas (quintiles 5 and 4 respectively).

The modal urban/rural classification for each population mixing category, for all years, was main urban area (Tables 6.13-6.15). This finding is hardly surprising since the majority of CAUs in New Zealand (54 percent) are assigned to this category. To examine urban/rural trends in more detail, the number of CAUs in each population mixing category was classified as either urban or rural, and then examined as a percentage of the total urban or total rural CAUs in New Zealand. In 1981, the patterns of population mixing were similar between urban and rural areas (Figure 6.28). The majority of urban areas were either LMHD (42 percent) or HMLD (33 percent). The corresponding figures for rural areas were 32 and 45 percent respectively. The lowest (LMLD) and highest (HMHD) population mixing areas only represented a small proportion of urban and rural areas in 1981. For example, only 10 percent of urban areas were classed as HMHD, and only 12 percent of rural areas were classed as HMHD. Some substantial changes had occurred in these values by 2001 (Figure 6.29). The majority of urban areas in 2001 were classed as HMHD (37 percent), and a large increase in the percentage of urban CAUs classed as LMLD, had also occurred. By 2001, rural areas were predominantly low population mixing areas; LMLD represented 40 percent and LMHD represented 42 percent of all rural areas in New Zealand. Only 11 percent of rural areas were classed as HMHD.

Table 6.13: Population mixing categories by population, deprivation and urban/rural variables, 1981

Population mixing category 1981	Mean population 1981	Mean population density 1991	Modal deprivation quintile 2001	Modal urban/rural category 2001
Low migration low diversity	1,806.63	968.21	4	main urban
Low migration high diversity	2,412.29	1,129.07	4/5	main urban
High migration low diversity	1,348.83	787.83	1	main urban
High migration high diversity	1,630.95	859.66	5	main urban

Table 6.14: Population mixing categories by population, deprivation and urban/rural variables, 1991

Population mixing category 1991	Mean population 1991	Mean population density 1991	Modal deprivation quintile 2001	Modal urban/rural category 2001
Low migration low diversity	1,742.41	781.99	1	main urban
Low migration high diversity	2,252.36	855.06	2	main urban
High migration low diversity	1,464.07	831.09	5	main urban
High migration high diversity	2,367.39	1,313.51	3	main urban

Table 6.15: Population mixing categories by population, deprivation and urban/rural variables, 2001

Population mixing category 2001	Mean population 2001	Mean population density 2001	Modal deprivation quintile 2001	Modal urban/rural category 2001
Low migration low diversity	1,682.65	745.11	1	main urban
Low migration high diversity	2,178.58	581.07	2	main urban
High migration low diversity	1,989.44	1,103.64	5	main urban
High migration high diversity	2,885.71	1,671.12	4	main urban

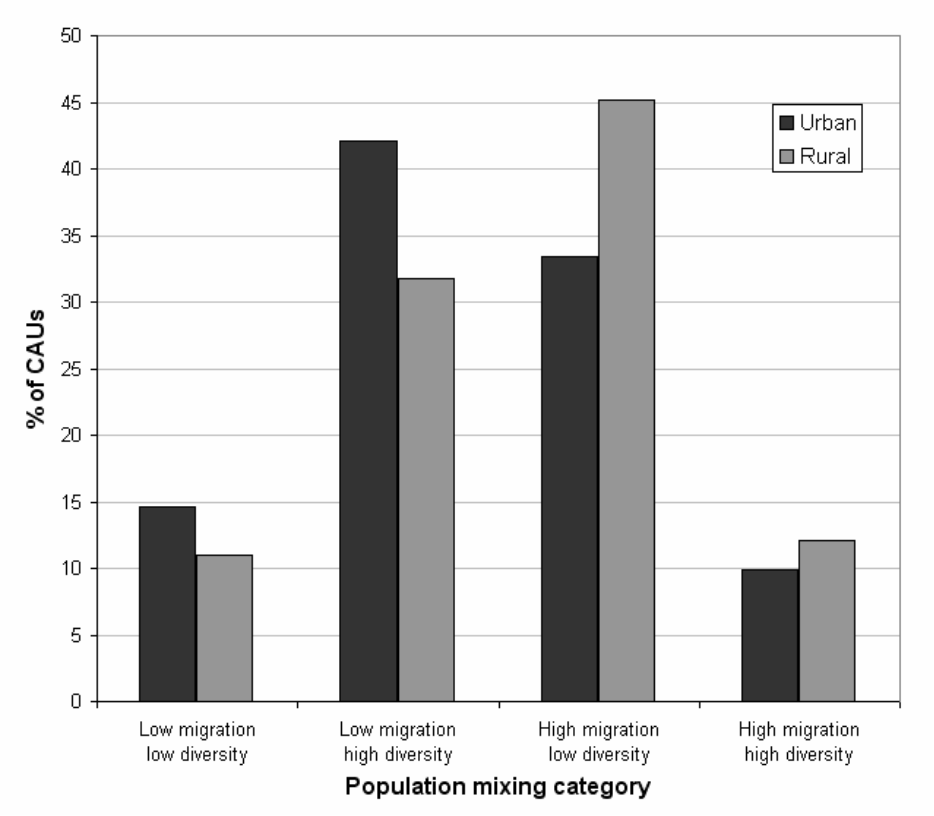

Figure 6.28: Population mixing category by urban/rural classification of CAUs, 1981

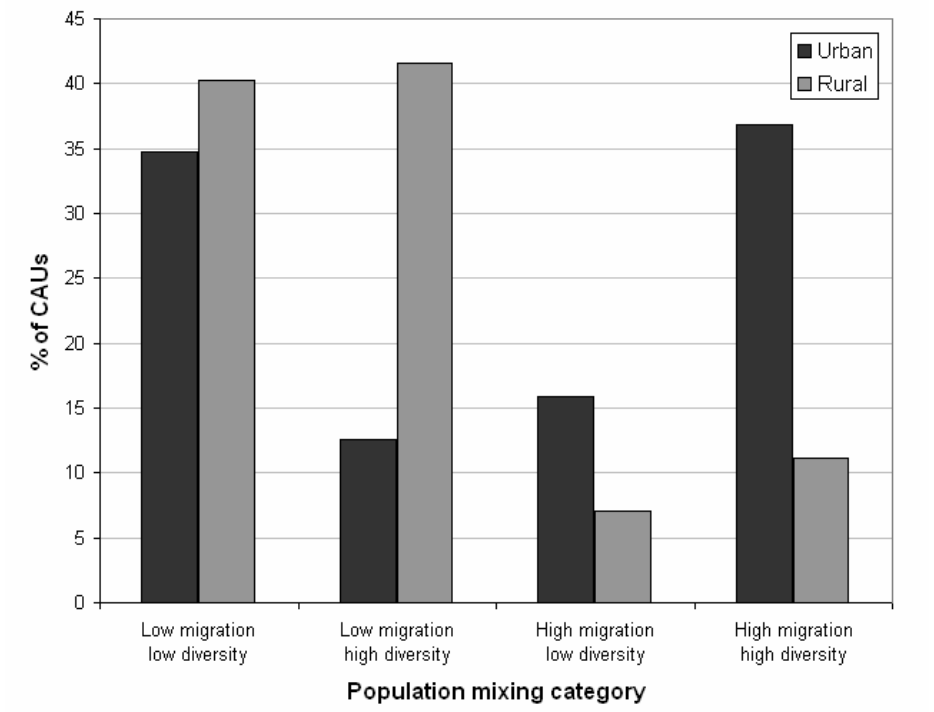

Figure 6.29: Population mixing category by urban/rural classification of CAUs, 2001

6.8 Change in population mixing categories

Since examining temporal changes in population mixing is a central part of the aims of this research, it was important to establish a systematic way of measuring the variation in population mixing categories over time. As a result, new population mixing change categories were devised. Every CAU in New Zealand was thus assigned to one of four new categories:

1. Decrease: areas which decreased in either the in-migration or migrant diversity category.
2. No change: areas which did not change in either category.
3. Slight increase: areas which increased in the migrant diversity category, but *not* the migration category.
4. Large increase: areas which increased in the in-migration category, and possibly increased in the migrant diversity category.

6.8.1 Description of the population mixing change categories 1981-2001

Category 2 was the most frequently occurring population mixing change category for every time period (Table 6.16). This category represented areas which had not changed in either in-migration or diversity category during the study period. Moreover, the importance of this category increased over time. For example, between 1981 and 1986, 44.76 percent of New Zealand CAUs were classed as category 2 areas, whereas between 1996 and 2001, this figure had increased to 70.63 percent. Category 1 areas, where either in-migration or migrant diversity levels decreased to below the national average, were the second most frequently occurring areas in New Zealand, for all time periods. Therefore the overwhelming majority of New Zealand CAUs did not witness substantial (from below to above average) increases in population mixing during the five year study periods. Between four and eight percent of CAUs increased to above average migrant diversity during any period (category 3), and between eight and 16 percent of CAUs increased to above average in-migration (category 4). The relative importance of category 4 areas, which should be the most important areas in terms of the population mixing hypothesis, decreased over time. For the period 1981-1986 category 4 areas represented 15.44 percent of the total CAUs in New Zealand. This figure decreased to 8.55 percent during the period 1991-1996 and rose slightly to 10.38 during the final five years (1996-2001).

Table 6.16: Population mixing change category by the percentage of total CAUs and year 1981-2001

Population mixing change category	Category Abbreviation	% of total CAUs			
		1981-86	1986-91	1991-96	1996-01
1. Decrease in either/both categories	Decrease	32.44	17.05	15.71	13.45
2. No change in either category	No change	44.76	66.16	69.39	70.63
3. Increase in diversity category	Slight increase	7.37	4.30	6.35	5.54
4. Increase in migration category	Large increase	15.44	12.48	8.55	10.38

6.8.2 Geography of the population mixing change categories 1981-2001

The patterns summarised in Table 6.16 are clearly visible when examining the geography of the population mixing change categories (Figure 6.30). In the first half of the study period (1981-1991) the majority of the North Island CAUs were categorised as 'no change' (category 2). More specifically, the region of Manawatu-Wanganui, southern Waikato and the majority of the Bay of Plenty region did not change substantially in population mixing during this time. Areas of decrease (category 1) were spatially contiguous in the north, east and west capes of the island, and also in the majority of the Hawkes Bay region, and the east coast of the Wellington region. Areas of slight and large increases in population mixing (categories 3 and 4) were generally noted in the main towns and cities, for example in central Auckland, Tauranga, Hamilton, Napier, Taupo and Wellington. In the South Island, the majority of the western half and northern section of the island decreased in population mixing category between 1981 and 1991. The southern part of the island mostly noted no change in population mixing category, as did large areas in Canterbury in the east. CAUs which increased in population mixing category (3 and 4) were located in Nelson, Blenheim, Fox Glacier, Franz Joseph Glacier, Invercargill, Dunedin, and in a zone around the city centre in Christchurch.

Comparing these patterns to those for the second half of the study period (1991-2001), the number of CAUs classed as no change (category 2) was clearly higher. By 1991-2001, even more of the North Island was characterised by no substantial change in population mixing. A few areas of decrease (category 1) were noted in the far north CAU, CAUs south of Auckland, and some areas in Hawkes Bay, eastern Taranaki and north of Wellington. A notable area of large increase (category 4) occurred in CAUs west of Lake Taupo and extended as far north as Te Awamutu. Between 1991 and 2001, most of the South Island CAUs were categorised as areas of no change (category 2). Exceptions included a few CAUs which decreased in population mixing category (category 1) located on the north-west coast. Moreover, slight increases in population mixing (category 3) occurred in CAUs on the south and south-east coast, in the Waimakariri District in Canterbury and in alternate CAUs in the north of the island. Large increases in population mixing (category 4) were

noted in parts of the central West Coast, Queenstown and Glenorchy, Stewart Island, and CAUs in the outskirts of Christchurch.

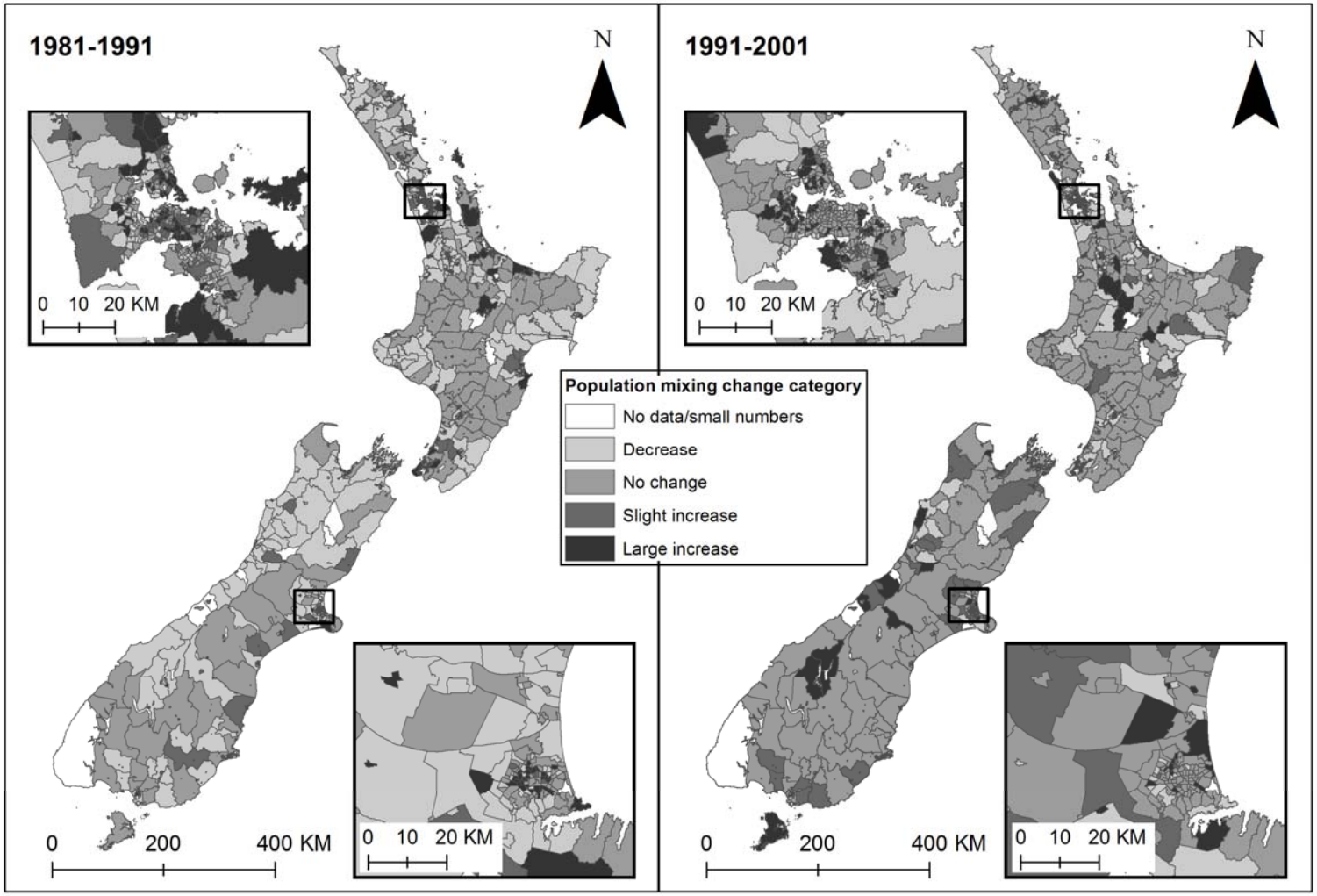

Figure 6.30: Population mixing change category by CAU with insets of Auckland in the North Island and Christchurch in the South Island, for 1981-1991 and 1991-2001

6.8.3 Contextual understanding of the population mixing change categories 1981-2001

During the period 1981-1991, the average population of New Zealand CAUs increased with increasing population mixing change category (Table 6.17). The highest population areas were also those which experienced a large increase in population mixing (category 4) between 1981 and 1991. A similar pattern was evident with the mean population density of CAUs, except the highest mean population density (1,357.86 people/km^2) occurred in category 3 (slight increase), rather than category 4, areas. There was no discernible trend in population mixing change by deprivation quintile for this period. The modal deprivation quintile in decrease areas was quintile 5 (most deprived), in no change areas was quintile 3, in slight increase areas was quintile 2, and was quintile 4 in large increase areas.

In contrast, the modal deprivation category increased with increasing population mixing change category between 1991 and 2001 (Table 6.18). The most frequently occurring deprivation quintile in category 1 (decrease) areas, was quintile 1 (most affluent) compared to quintile 5 (most deprived) in population mixing change category 4 (large increase) areas. Therefore the most deprived areas of the country tended to be those which increased to above average in-migration between 1991 and 2001. There was no clear trend in mean population and mean population density across the population mixing change categories for this time period. The highest average population of 2,565.28 people per CAU occurred in category 4 (large increase) areas, as did the highest average population density (1,397.82 people/km^2). However, the next highest values for these measures were noted in category 2 (no change) areas.

For both time periods, the modal urban/rural classification for each population mixing change category was main urban area. However, when examining the population mixing change categories by the percentage of CAUs they accounted for in either urban or rural areas of New Zealand, interesting patterns in rural areas were also evident (Figures 6.31 and 6.32). For example, between 1981 and 1991, the majority of rural CAUs either decreased (51 percent) or witnessed no substantial change (34 percent) in population mixing (Figure 6.31). While most of rural New Zealand was still characterised by one of these two categories during the period 1991-2001, only 22 percent of rural CAUs decreased in population mixing compared to 61 percent of rural CAUs which did not change greatly over the period (Figure 6.32). The population mixing change categories were slightly more evenly spread among the urban CAUs of New Zealand between 1981 and 1991, but the decrease and no change categories still dominated, with 36 percent of urban areas classed as no change (Figure 6.31). By the second half of the study period, this figure had increased to 67 percent (Figure 6.32).

Table 6.17: Population mixing change categories by population, deprivation and urban/rural variables, 1981-1991

Population mixing change category 1981-1991	Mean population 1991	Mean population density 1991 (km^2)	Modal deprivation quintile 2001	Modal urban/rural category 2001
Decrease	1,696.54	698.74	5	main urban
No change	2,039.60	934.40	3	main urban
Slight increase	2,223.52	1,357.86	2	main urban
Large increase	2,317.27	1,259.93	4	main urban

Table 6.18: Population mixing change categories by population, deprivation and urban/rural variables, 1991-2001

Population mixing change category 1991-2001	Mean population 2001	Mean population density 2001 (km^2)	Modal deprivation quintile 2001	Modal urban/rural category 2001
Decrease	1,804.86	742.33	1	main urban
No change	2,275.89	1,075.59	3	main urban
Slight increase	1,708.68	780.03	5	main urban
Large increase	2,565.28	1,397.82	5	main urban

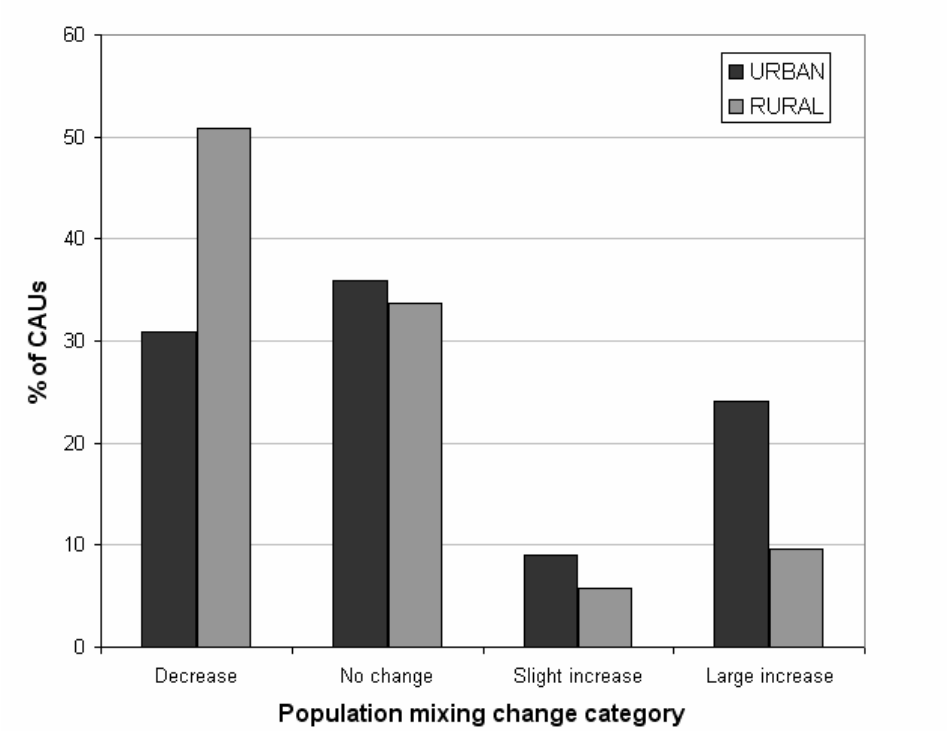

Figure 6.31: Population mixing change category by urban/rural classification of CAUs, 1981-1991

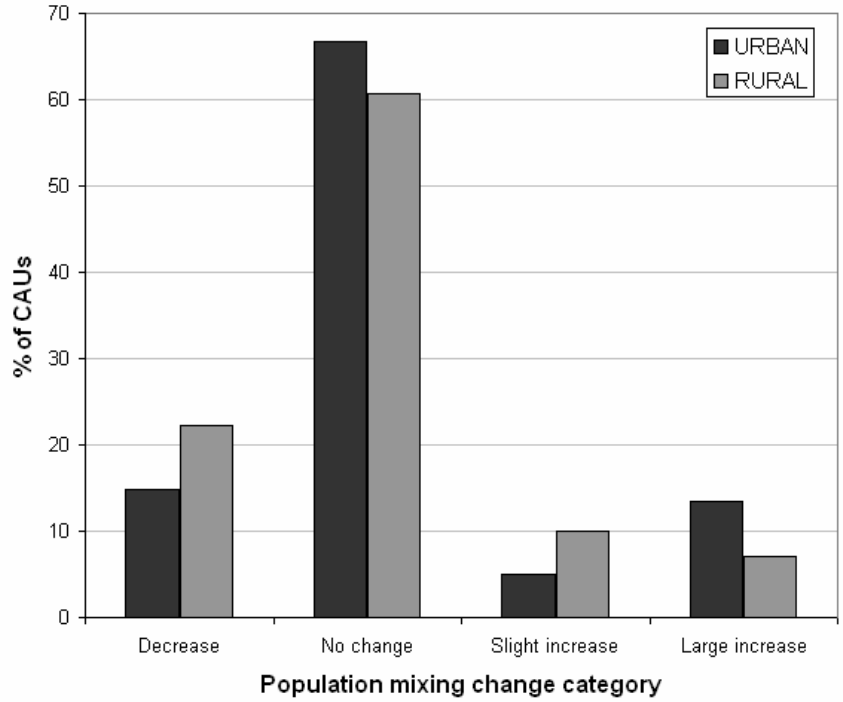

Figure 6.32: Population mixing change category by urban/rural classification of CAUs, 1991-2001

6.9 Conclusion

A recurring theme throughout this chapter has been the temporal variability of population movements in New Zealand between 1981 and 2001. Over this 20 year period there has been a general increase in population mixing in New Zealand. However, the examination of five year periods revealed considerable shorter-term fluctuations. Furthermore, temporal trends varied depending upon the population mixing measure under consideration. A second important theme is that there were large spatial variations in the population mixing measures across the country. For example, total population growth tended to be highest in the north and north-east of the North Island and the north-east and central south of the South Island. At a more detailed geographical scale (CAU-level), variation within these general trends was also revealed. Additionally, these geographical patterns changed over time. Finally, this chapter has also shown that population mixing varies according to a number of contextual factors. There was a strong social gradient noted for population change between 1981 and 2001, with the highest population growth occurring in the most affluent CAUs for all time periods. Interestingly, the opposite relationship was noted between deprivation quintile and the majority of the other population mixing measures. Patterns by the urban/rural nature of CAUs were more complex, and depended upon the population mixing measure and time period.

While many important aspects of population movements within New Zealand have been described during the course of this chapter, it is not the purpose of this thesis to offer an explanation for such movements. What is central to the aims of this research is the possible importance of the population mixing patterns detailed here to childhood disease causation; in particular, the small area incidence of childhood ALL and type 1 diabetes. The next two chapters of the thesis display the results of the analyses which test for associations between population mixing and childhood ALL (Chapter 7) and type 1 diabetes (Chapter 8).

Chapter 7: Population mixing and the geographical epidemiology of childhood acute lymphoblastic leukaemia in New Zealand

7.1 Introduction

The previous chapter demonstrated that movements and mixing of people in New Zealand varied considerably both spatially and temporally between 1981 and 2001, particularly by urban/rural classification and areal deprivation. The aim of this chapter is to examine whether these trends in population mixing help to explain the variation in incidence of acute lymphoblastic leukaemia (ALL) in children aged 0 to 14 years. First, the chapter details the geographical epidemiology of this disease in New Zealand. Second, a number of univariate and multivariate methods are utilised to examine the relationship between ALL and population mixing.

7.2 Descriptive patterns

7.2.1 Individual-level

Previous international studies describe how ALL varies by age at diagnosis, sex, and ethnic origin (Greaves, 1997, Little, 1999, McNally and Parker, 2006). Between 1980 and 2004 (inclusive) a total of 781 new cases of ALL were registered in children aged 0-14 years in New Zealand. Of these new cases, 456 were male (58.4%) and 325 were female (41.6%). The count of ALL cases peaked among those aged 0-4 years at diagnosis and accounted for 55.8% of the total cases. The mean age at diagnosis was slightly higher for males (5.27) than females (5.07). The median age at diagnosis was 4 years for both females and males, whereas the modal years were 2 and 3 respectively (Table 7.1). These figures show that on average, the majority of females in this cohort developed ALL around one year earlier than the males.

Table 7.1: ALL descriptive statistics by age and sex

Descriptive Statistics	Male	Female	Total
Count	456	325	781
% of total	58.4	41.6	100
Mean age at diagnosis	5.27	5.07	5.19
Modal age at diagnosis	3	2	2
Median age at diagnosis	4	4	4
Standard deviation of age at diagnosis	3.70	3.61	3.67

The total number of ALL registrations over this period was consistently higher in males of all ages, and the sex difference was particularly marked among those aged 2 to 4 years (Figure 7.1). The biggest difference between males and females was observed at age 3 when 79 cases were recorded in males compared with only 47 cases in females. However, these figures do not account for the denominator population. When the ALL cases were grouped into five year age groups and divided by the age- and sex-specific populations at risk, males had higher rates of ALL across all three age groups (Figure 7.2). This trend was especially pronounced in the 0-4 year age group, where nearly 180 ALL cases per 100,000 were observed in males, compared to 135 cases per 100,000 in females.

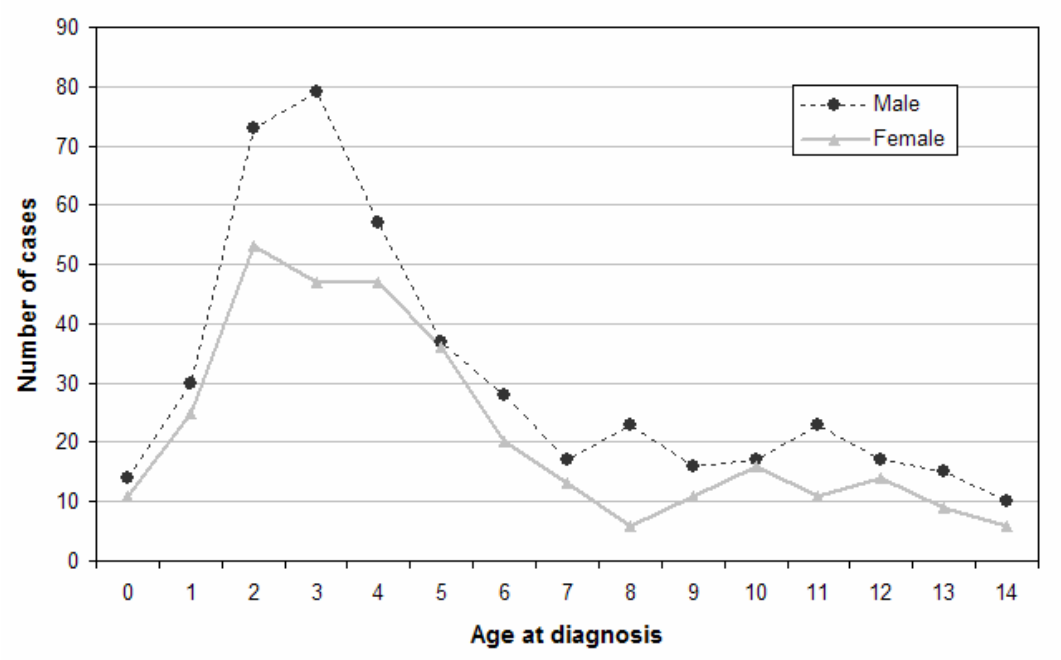

Figure 7.1: Number of ALL cases by age at diagnosis and sex, 1980-2004

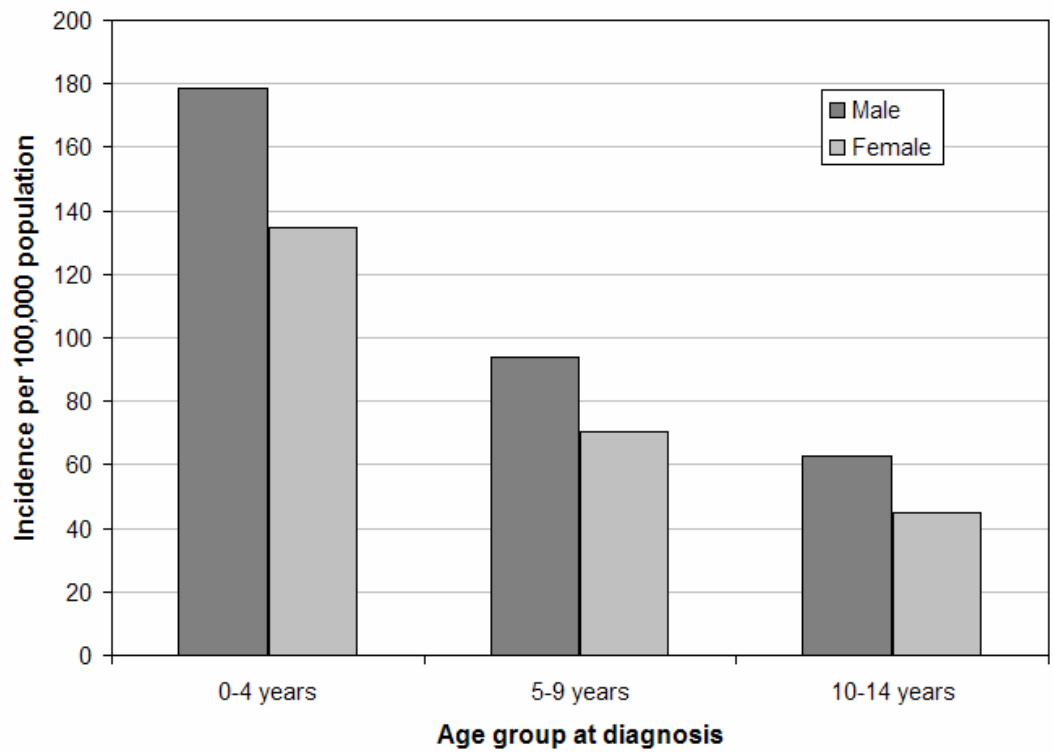

Figure 7.2: ALL incidence per 100,000 population by age group and sex, 1980-2004

There were also important differences in the incidence of ALL between the main ethnic groups. Of the total ALL cases registered between 1980 and 2004, 75.42 percent were of European ethnicity (Table 7.2). An excess of European cases was also observed after accounting for the population at risk in each of the four largest ethnic groups in New Zealand (Figure 7.3). Incidence of ALL was highest among European children, with around 95 new cases diagnosed per 100,000 population between 1980 and 2004 compared to 64 cases in Pacific children, 55 cases in Asian children and 50 cases in Māori children. Due to the relatively small number of ALL cases by age group in Māori, Pacific and Asian children, age-specific incidence rates were not calculated. However, peaks in the total number of cases by all ethnic groups were noted in the 0-4 year age group (Table 7.2)

Table 7.2: ALL counts by ethnic group, age group at diagnosis and sex

Ethnic Group	Sex	0-4 years	5-9 years	10-14 years	All ages	% of total cases
European	Male	188	93	61	342	43.79
	Female	137	65	45	247	31.63
	Total	325	158	106	589	**75.42**
Māori	Male	40	15	10	65	8.32
	Female	21	7	5	33	4.23
	Total	61	22	15	98	**12.55**
Pacific	Male	14	6	8	28	3.59
	Female	18	10	2	30	3.84
	Total	32	16	10	58	**7.43**
Asian	Male	9	6	2	17	2.18
	Female	6	4	4	14	1.79
	Total	15	10	6	31	**3.97**
Unknown	Male	1	0	1	2	0.26
	Female	1	0	0	1	0.13
	Total	2	0	1	3	**0.38**
Other	Male	1	1	0	2	0.26
	Female	0	0	0	0	0.00
	Total	1	1	0	2	**0.26**
Total		436	207	138	781	**100.00**

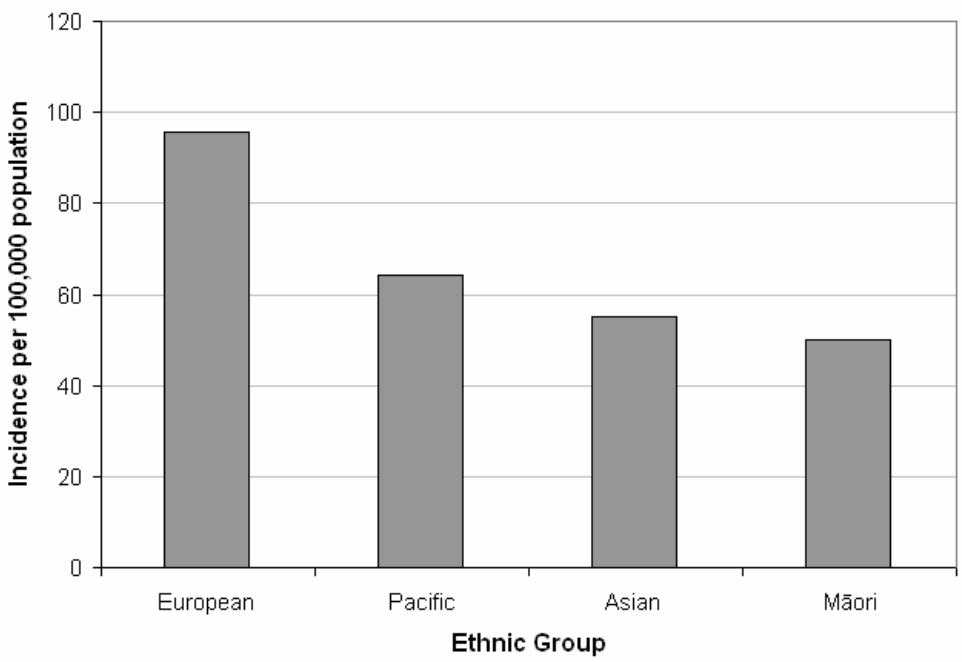

Figure 7.3: ALL incidence by ethnicity 1980-2004

Examination of temporal trends (Figure 7.4) revealed a gradual increase in the incidence of ALL over the whole study period. In 1980, 2.6 new cases of ALL were diagnosed per 100,000 population compared to 3.3 new cases in 2004. After 1985, the incidence of ALL only fell below 3 cases per 100,000 once (in 1998). Three peaks in incidence occurred: in 1983 (4.4 new cases), 1987 (4.9 new cases) and 2000 (5.6 new cases per 100,000 population). Between 1988 and 1996,

the incidence of childhood ALL remained relatively stable at around four new cases per 100,000 population per year. According to the linear regression equation for these data, the average increase in incidence was 0.04 cases per 100,000 population per year.

As a result of the small number of cases diagnosed by year and age group, incidence rates were calculated using three year moving averages. ALL incidence per 100,000 population was considerably greater in children aged 0-4 years at diagnosis compared to the other age groups, for every year (Figure 7.5). The lowest incidence for this age group was observed in 1985 (4.5 cases per 100,000) and the highest in 2000 (8.9 cases per 100,000). Thus, the incidence of ALL in 0-4 year olds increased gradually over time and remained consistently above 5 cases per 100,000 per year after 1985. Incidence in the 5-9 year age group was higher than incidence in the 10-14 year age group with the exception of two years (1985 and 1986). After 1987, ALL in 5-9 year olds remained reasonably stable at just over 3 new cases per 100,000 per year. Incidence in the oldest age group (10-14 years) fluctuated between 1 and 3 new cases per 100,000 during the study period, with a gradual increase noted since 1991.

Male ALL incidence was generally higher than female ALL incidence in every year except for 1989 and 2003 (Figure 7.6). Both male and female incidence rates increased over the study period, with the lowest male incidence noted in 1985 (3.3 cases per 100,000) and the highest male incidence noted in 2000 (6.5 cases per 100,000). Female incidence of ALL peaked in 1989 (4.2 cases per 100,000), but then increased gradually after 1999. In Figures 7.4, 7.5 and 7.6, incidence of ALL decreased after 2000 (except in females). However, this finding may be an artefact of incomplete data collection in the most recent years.

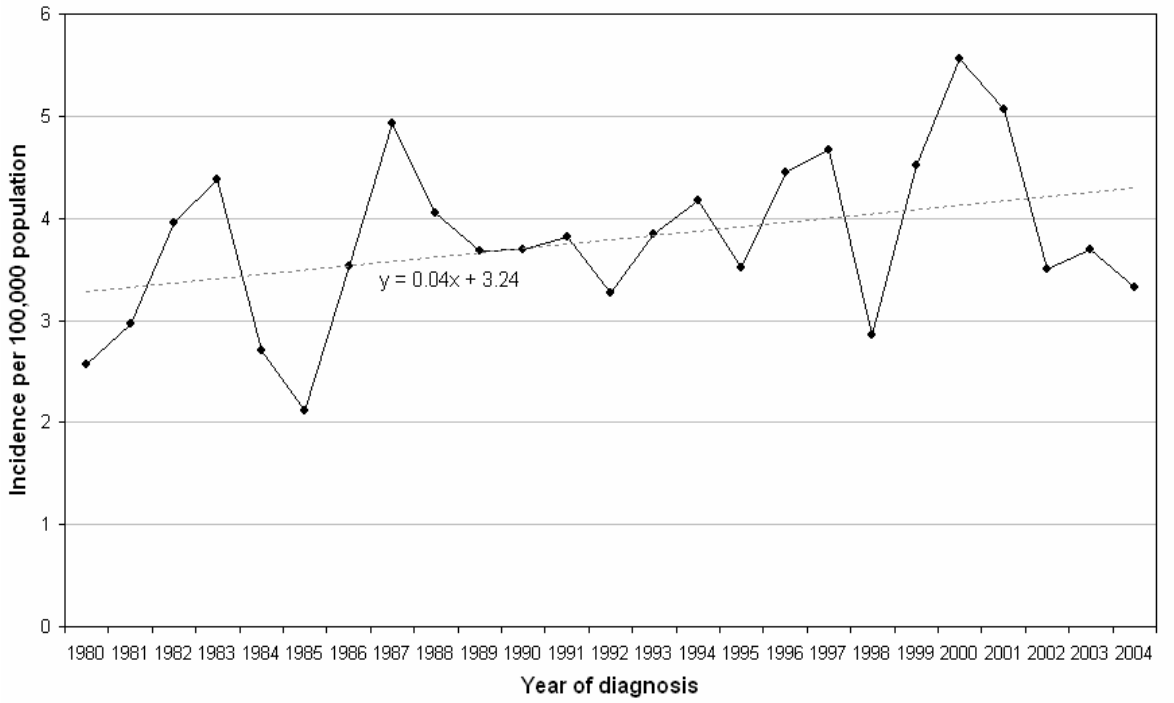

Figure 7.4: ALL incidence by year of diagnosis

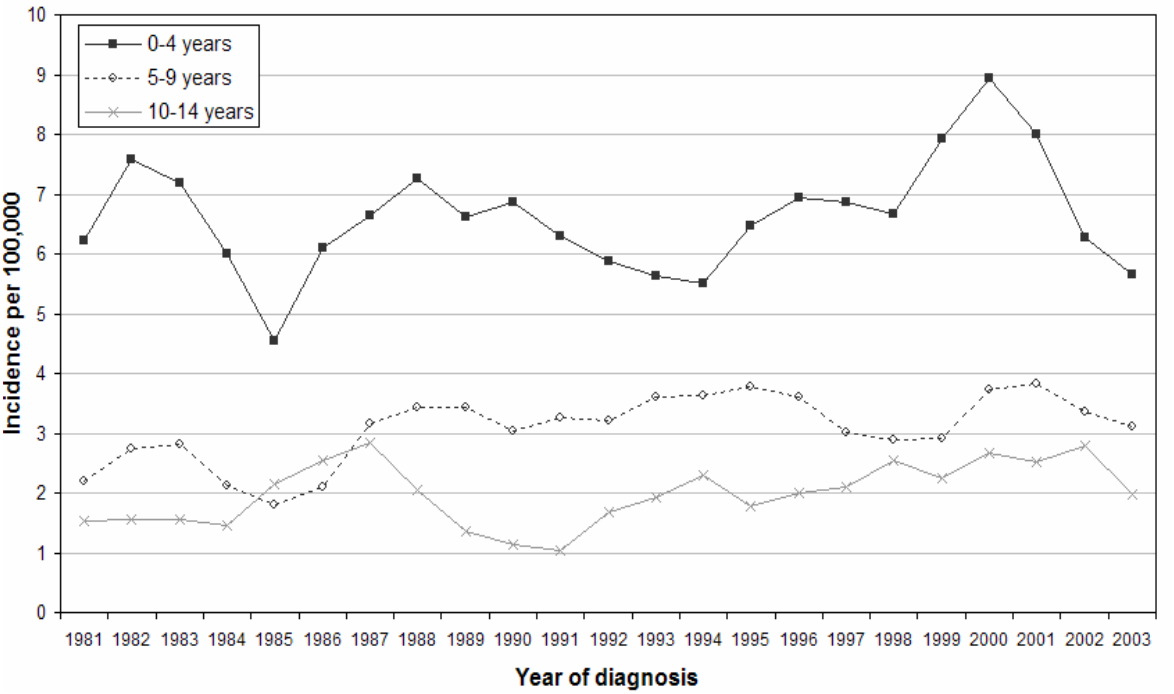

Figure 7.5: ALL incidence by age group and year of diagnosis (3 year moving averages)

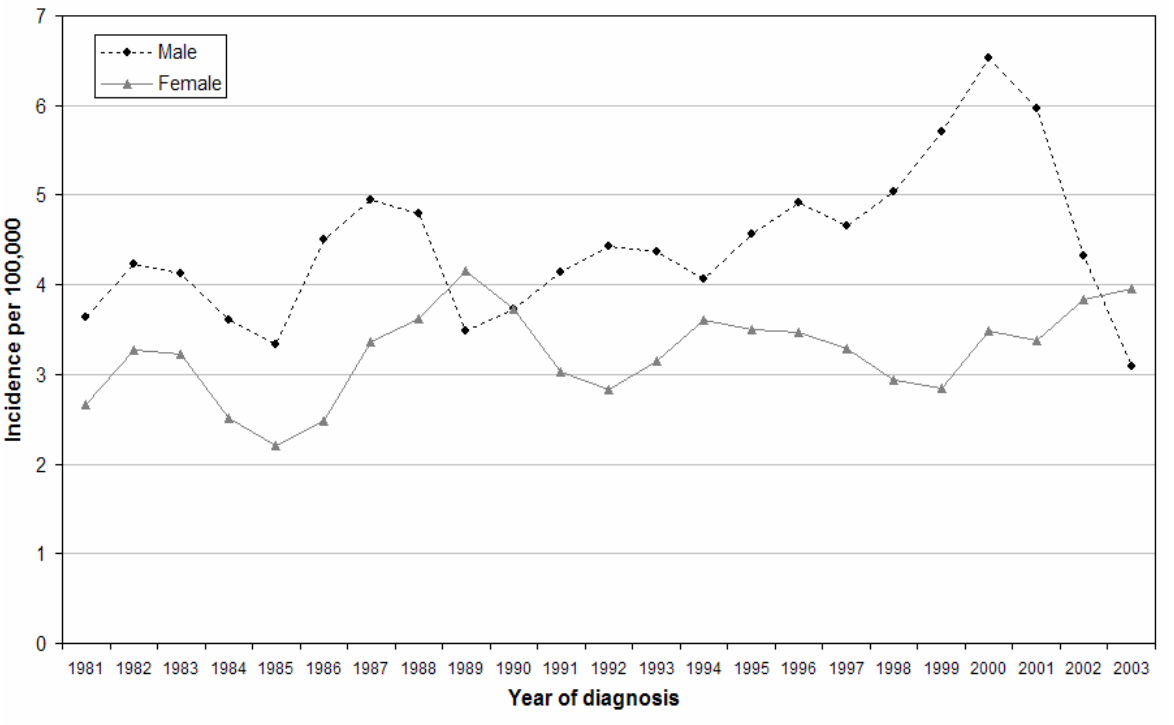

Figure 7.6: ALL incidence by sex and year of diagnosis (3 year moving averages)

7.2.2 Area-level

As well as describing how childhood ALL varies by individual-level variables such as age, sex, ethnicity and year of registration, it is also useful to examine the importance of any area-level influences. Trends in standardised incidence ratios (SIRs) were examined by deprivation deciles. SIRs of 100 indicate that the number of ALL cases observed in the deprivation decile is equal to the number of cases expected in the national population. SIRs which are greater than 100 indicate that more ALL cases occurred than expected, and SIRs which are less than 100 indicate that fewer ALL cases occurred than expected.

There was no clear relationship between ALL and deprivation decile (Figure 7.7 and Table 7.3). The decile with the highest SIR (132.82, CI = 100.47-165.18) was decile 3 (at the least deprived end of the spectrum) suggesting that those living in areas classed as deprivation decile 3 experienced an excess of ALL cases between 1980 and 2004 when compared to national rates. However, the SIRs of the least deprived deciles (1 and 2), were below 100 and suggest fewer than expected cases in these areas. Deciles 9 and 10 (at the most deprived end of the spectrum) also had SIRs below 100 (95.51 and 76.50 respectively), suggesting that the most deprived areas in New Zealand had fewer ALL cases than expected. However, the SIR for Decile 8 was greater than 100.

The widths of the confidence intervals for these SIRs (shown by the bars on the graph in Figure 7.7) were wide, suggesting that the SIR values were relatively unstable. This finding could be due to the SIRs being based on a small number of cases. Consequently, SIRs were also calculated by deprivation quintiles to reduce the error margin of these values. The trend in SIRs across deprivation quintiles was also non-linear (Table 7.4). Quintile 2 had the highest SIR of 111.45 (CI = 92.81-130.10), whereas quintile 1 had a SIR of just below 100. Quintile 5 had the lowest SIR of 85.12 (CI = 73.4-96.75) suggesting that ALL rates were significantly lower in the most deprived fifth of New Zealand CAUs (Table 7.4). The quintile and decile-level analysis both found that significantly fewer cases of ALL were witnessed in the most deprived areas of New Zealand over the study period.

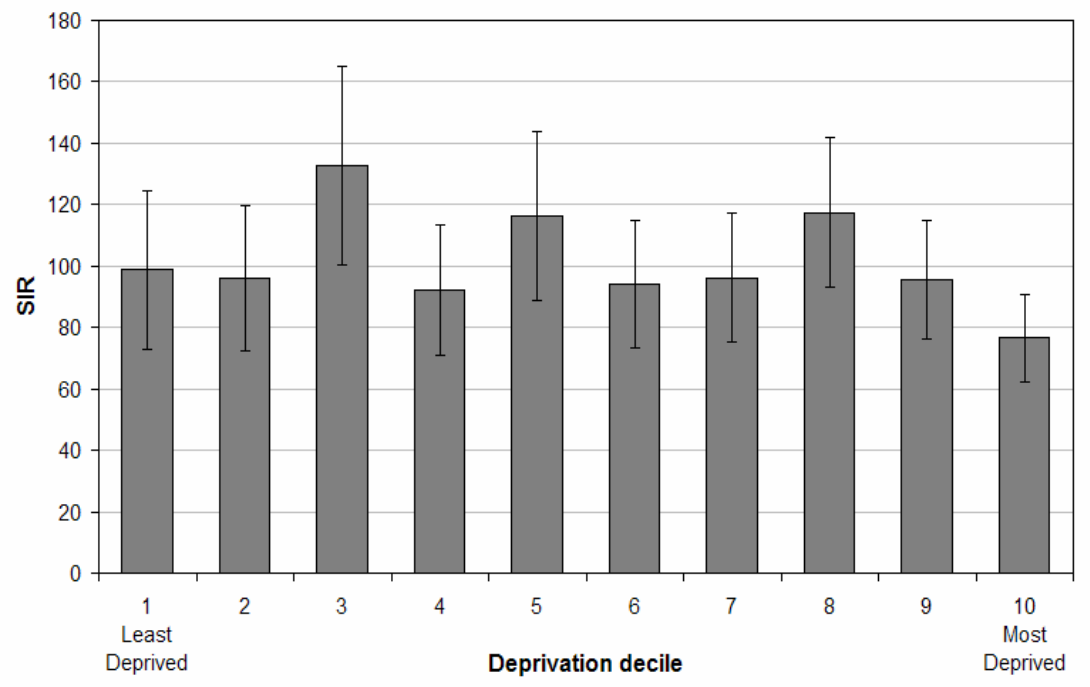

Figure 7.7: ALL SIRs by area deprivation decile, 1980-2004

Table 7.3: ALL SIRs, chi-square values and CIs by deprivation decile, 1980-2004

Deprivation Decile	Observed	Expected*	SIR	Lower CI	Upper CI	Chi-square	Significant
1 (Low)	55	55.71	98.73	72.80	124.66	0.01	NO
2	62	64.54	96.06	72.62	119.50	0.10	NO
3	86	64.75	132.82	100.47	165.18	6.98	YES
4	67	72.53	92.38	71.12	113.64	0.42	NO
5	79	67.94	116.28	88.63	143.93	1.80	NO
6	74	78.68	94.05	73.27	114.83	0.28	NO
7	78	81.13	96.14	75.22	117.06	0.12	NO
8	105	89.40	117.46	93.11	141.80	2.72	NO
9	89	93.18	95.51	76.12	114.91	0.19	NO
10 (High)	86	112.42	76.50	62.36	90.64	6.21	YES

*Expected cases = sum of (age-specific population in each decile x age-specific incidence rate)

Table 7.4: ALL SIRs, chi-square values and CIs by deprivation quintile, 1980-2004

Deprivation Quintile	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
1 (Low)	117	120.25	97.30	79.91	114.69	0.09	NO
2	153	137.28	111.45	92.81	130.10	1.80	NO
3	153	146.62	104.35	87.46	121.24	0.28	NO
4	183	170.52	107.32	91.21	123.42	0.91	NO
5 (High)	175	205.60	85.12	73.48	96.75	4.56	YES

A further area-level variable thought to influence childhood ALL incidence is the degree of area rurality (e.g. Adelman et al., 2005, Kinlen and Petridou, 1995). SIRs were thus calculated for seven categories of area across the urban/rural continuum in New Zealand (Table 7.5). Lower ratios were noted in most of the rural categories but there was no dose-response relationship along the urban/rural spectrum. The highest SIR (118.05) occurred in rural areas with a high urban influence. In contrast, all of the other rural areas (with lower degrees of urban interaction) had SIRs below 100, revealing that expected cases exceeded observed cases in these areas. The two most urban categories witnessed SIRs greater than 100, suggesting that an excess of ALL cases occurred in these areas over the 25 year period. For example, main urban areas were expected to have around 522 cases of ALL over the study period, but actually recorded 546 cases resulting in a SIR of 104.61 (CI = 95.64-149.39).

Table 7.5: ALL SIRs, chi-square values and CIs by urban/rural classification, 1980-2004

Urban/Rural Category	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
Main urban area	546	521.93	104.61	95.64	113.59	1.11	NO
Satellite urban community	27	24.39	110.71	66.77	154.65	0.28	NO
Independent urban community	95	103.84	91.49	73.89	109.09	0.75	NO
Rural high urban infl.	25	21.18	118.05	67.77	168.33	0.69	NO
Rural moderate urban infl.	24	26.66	90.03	55.85	124.21	0.26	NO
Rural low urban infl.	52	67.18	77.40	58.89	95.91	3.43	NO
Highly rural/remote	12	15.61	76.88	38.74	115.02	0.83	NO

Due to the small number of observed cases in some of the urban/rural classes and the large error margins of the resulting SIRs, the original seven categories were aggregated into two main groups: predominantly urban areas (which included main urban areas, satellite urban communities and independent urban communities) and predominantly rural areas (all rural areas with varying levels of urban influence). A striking trend of higher rates in urban areas compared to rural areas was apparent. Urban areas had a SIR of 102.74 (CI = 94.85-110.64) compared to 86.51 (CI = 71.67-101.34) for rural areas (Table 7.6). However, the confidence intervals for these SIRs overlap.

Table 7.6: ALL SIRs, chi-square values and CIs by predominantly urban and rural areas, 1980-2004

Urban/Rural Category	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
Predominantly urban	668	650.16	102.74	94.85	110.64	0.49	NO
Predominantly rural	113	130.63	86.51	71.67	101.34	2.38	NO

7.3 Geographical distribution

The area-level analyses above showed that various aspects of the geographical context in which people live were associated with ALL incidence between 1980 and 2004. These analyses grouped together areas with similar features (such as high deprivation or high levels of rurality), and then compared SIRs of ALL across the categories. A way of detecting other possible place-level effects is to map disease rates. This method allows the identification of specific geographical areas with higher or lower than expected rates of disease and can help to generate hypotheses about potential environmental causes of the illness.

7.3.1 Geographical distribution of ALL SIRs

SIRs were thus calculated and mapped by geographical areas. The analysis was undertaken at two scales; at the territorial authority (TA) and census area unit (CAU)-level. New Zealand is divided in to 74 TA areas comprising of 15 cities and 59 districts (Figure 7.8). In 2001, the average population of a TA was approximately 50,500. The majority of TAs (47) recorded SIRs of between 1 and 100. A total of 27 out of the 74 TAs had SIRs over 100, 8 of which were cities with a population of over 50,000. At the TA-level, there was no clear spatial pattern to ALL incidence across New Zealand. Areas of high incidence were interspersed with areas of low incidence. It should be noted that the ratios calculated at the TA-level were quite unstable, as shown by their wide confidence intervals which frequently included 100. In fact, of the TAs with SIRs greater than 100, only two were significantly over 100 according to their confidence intervals: Christchurch City (SIR = 140.61, CI = 103.84-177.38) and Wellington City (SIR = 176.72, CI = 130.74-233.63). Of the TAs with SIRs under 100, only two were significantly so: Timaru District (SIR = 33.56, CI = 6.94-98.11) and Manukau City (SIR = 72.60, CI = 52.75-97.47). The areas of significantly high and low incidence ratios were not geographically close (Figure 7.9).

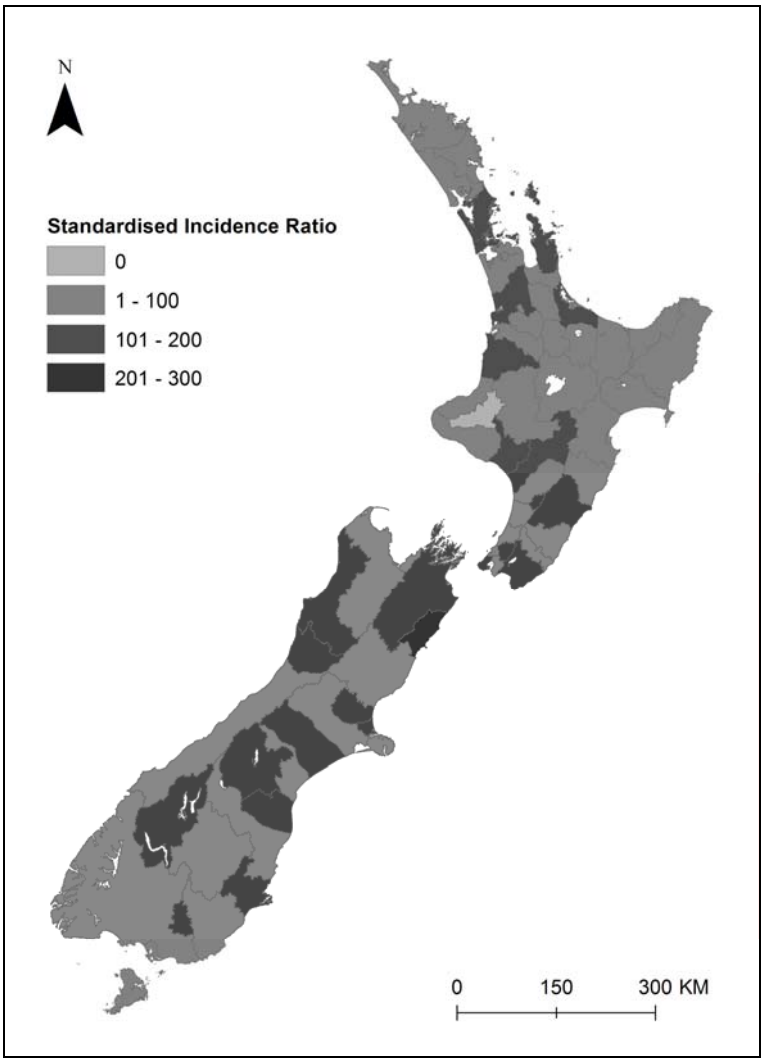

Figure 7.8: ALL SIRs by territorial authority, 1980-2004

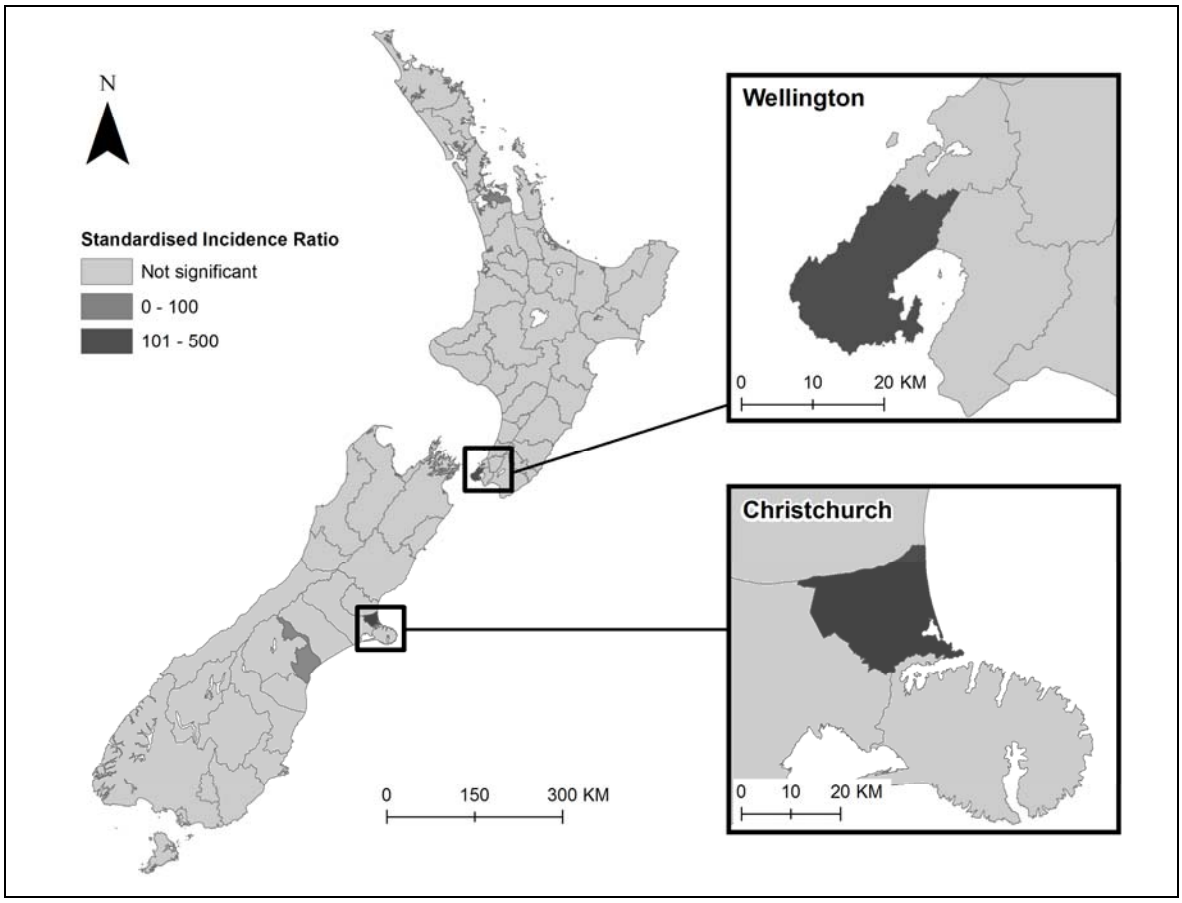

Figure 7.9: Significant ALL SIRs by territorial authority, 1980-2004

Although analysis at the TA-level provides an overview of the geography of ALL in New Zealand, it is likely that there was substantial variation at a more local scale. Consequently, SIRs for children with ALL (1980-2004) were also calculated and mapped at the census area unit (CAU)-level (Figure 7.10). The spatial pattern of ALL SIRs at the CAU-level was far more distinct than at the TA-level of analysis. The North Island had a concentration of areas with high incidence ratios, especially to the north and west of the island. Most of the west coast of the North Island, with the exception of the East Cape and Napier areas, had SIRs of zero. The majority of CAUs with SIRs of zero occurred in the southern half of the South Island. The northern half of the South Island also had many areas with SIRs of zero, but these were dispersed with CAUs with SIRs of over 100. The preponderance of SIRs of zero reflected the large number of CAUs in which no cases of ALL were observed over the twenty-five year period and highlights the rare nature of this disease.

In all of the main urban centres there was a patchwork of high and low SIRs (Figures 7.11 and 7.12). The majority of the highest SIRs (over 1000) were located within these areas, often in the city centre. Exceptions included the rural CAUs of Taramakau, Ahaurau and Mokihinui on the

North West coast of the South Island, and Rahotu on the West Cape of the North Island. The highest SIR recorded was 10,434.35 (CI = 313.03-58,119.30) for a CAU in central Whangarei in the North Island. However, this SIR was based on only one observed case. The highest SIR recorded in the South Island was in central Christchurch (SIR = 5,920.97, CI = 710.52-21,374.72) where 2 cases of ALL were observed compared with 0.03 expected cases. Both of these SIRs were statistically significant according to their chi-square values and confidence intervals (Table 7.7). As with the TA-level analysis, the wide range of the SIR confidence intervals indicated that the SIRs were unstable and could be a lot higher or lower than the calculated SIR suggests.

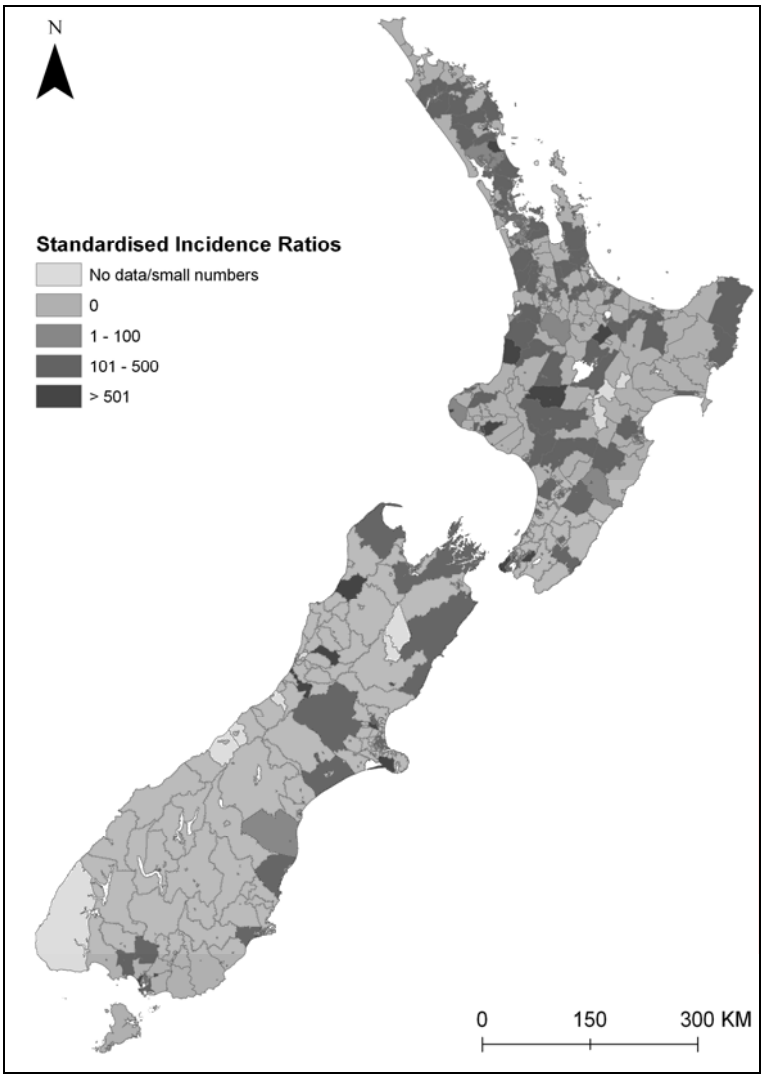

Figure 7.10: ALL SIRs by census area unit, 1980-2004

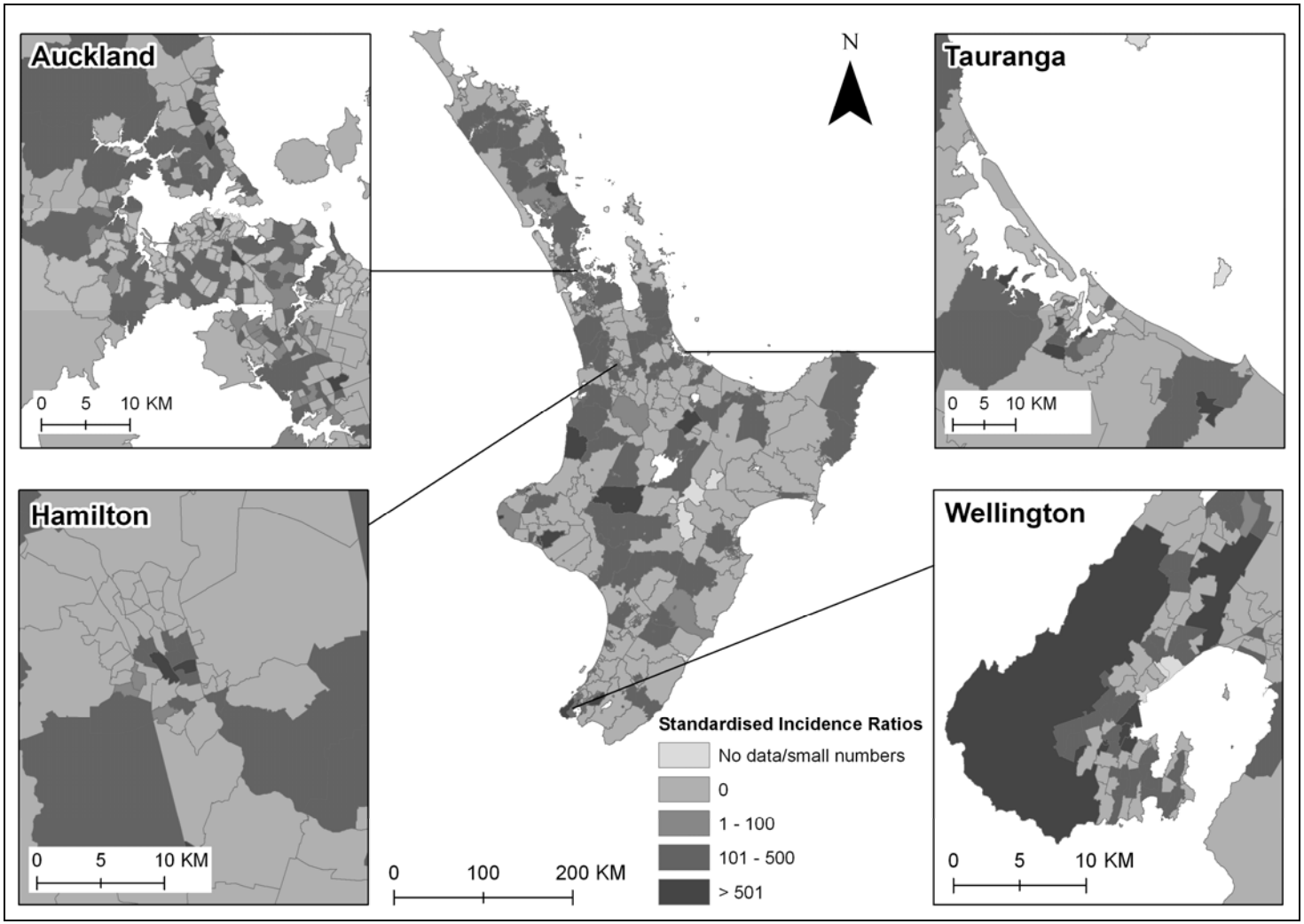

Figure 7.11: ALL SIRs by census area unit; North Island 1980-2004

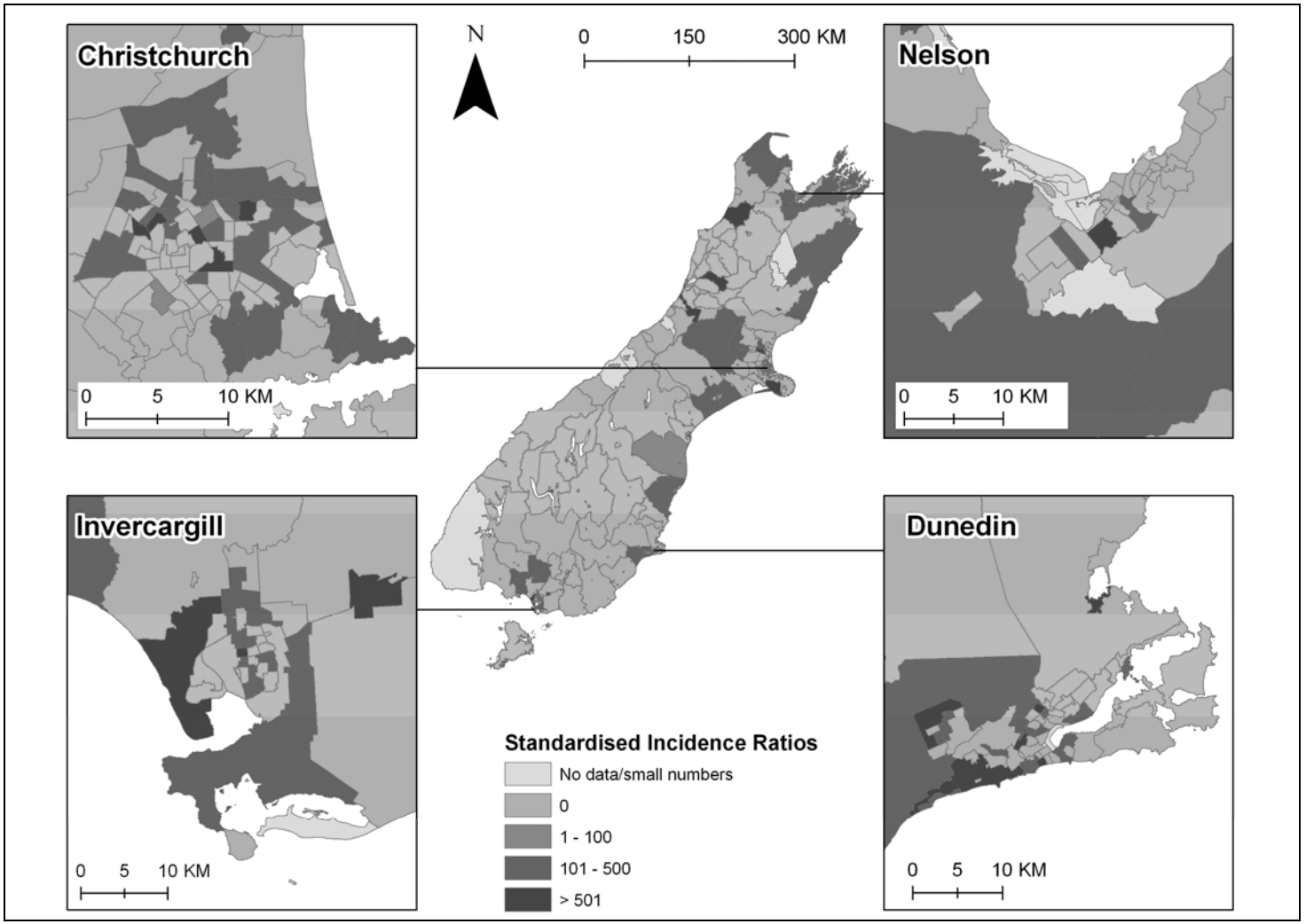

Figure 7.12: ALL SIRs by census area unit; South Island 1980-2004

Table 7.7: ALL SIRs over 1,000 by CAU 1980-2004

CAU	TA	SIR	Observed	Expected	LCI	UCI
Whangarei Central	Whangarei District	10434.35	1	0.01	313.03	58119.30
Cathedral Square	Christchurch City	5920.97	2	0.03	710.52	21374.72
Willis St	Wellington City	4269.09	1	0.02	128.07	23778.81
Taramakau	Westland District	4133.17	1	0.02	124.00	23021.77
Ngawhatu	Nelson City	3644.62	1	0.03	109.34	20300.51
Auckland Central W	Auckland City	3385.51	1	0.03	101.57	18857.30
Otakaro Park	Invercargill City	3241.79	2	0.06	389.02	11702.88
Pyes Pa	Tauranga District	2301.87	1	0.04	69.06	12821.39
Wyllies Crossing	Dunedin City	1750.66	1	0.06	52.52	9751.17
Taitville	Wellington City	1740.53	1	0.06	52.22	9694.77
Grenada	Wellington City	1666.90	1	0.06	50.01	9284.64
Kaikorai Lagoon	Dunedin City	1664.65	1	0.06	49.94	9272.11
Woodlands	Southland District	1560.40	1	0.06	46.81	8691.43
Makara-Ohariu	Wellington City	1534.29	2	0.13	184.12	5538.80
Ahaura	Grey District	1505.73	1	0.07	45.17	8386.90
Mokihinui	Buller District	1405.84	1	0.07	42.18	7830.51
Central Gore	Gore District	1322.55	2	0.15	158.71	4774.39
Mitchelltown	Wellington City	1263.02	1	0.08	37.89	7035.03
Kelvin Heights	Queenstown-Lakes District	1101.88	1	0.09	33.06	6137.50
Rahotu	South Taranaki District	1029.93	1	0.10	30.90	5736.70
Ohaupo	Waipa District	1029.07	1	0.10	30.87	5731.92
Pinehill	North Shore City	1021.40	2	0.20	122.57	3687.25
Runciman	Franklin District	1018.62	1	0.10	30.56	5673.69

7.3.2 Geographical distribution of ALL Poisson probabilities

SIRs are useful for examining how a disease varies by geographical area. However, when small numbers are involved, SIR values may be unreliable. Consequently, Poisson probabilities were also calculated to identify areas of higher or lower than expected incidence, compared to the Poisson distribution. Areas with Poisson probabilities under 0.05 represent places where there were significantly more cases of ALL than would be predicted by the Poisson distribution (at the 95 percent confidence level).

The majority of CAUs in New Zealand had ALL Poisson probability values which were greater than 0.05, suggesting that most CAUs had a similar number of cases to that expected from a Poisson distribution (Figure 7.13). The few areas with significant excesses of ALL cases were mainly found in and around the main urban areas of the country. For example, areas which had probabilities of 0.01 were all in the main urban centres of Whangarei, Auckland, Tauranga, Whakatane, Palmerston North, Christchurch, Gore and Invercargill, with the exception of one CAU (Makara-Ohariu) on the outskirts of Wellington classed as a rural area with a high urban influence. Areas where the Poisson probabilities were between 0.01 and 0.05 included Hokianga

North, Auckland, Hikuai, Hamilton, Tauranga, Rotorua, Wanganui, Palmerston North, Wellington, Nelson, Taramakau, Christchurch, Oamaru, and Dunedin (Figure 7.13).

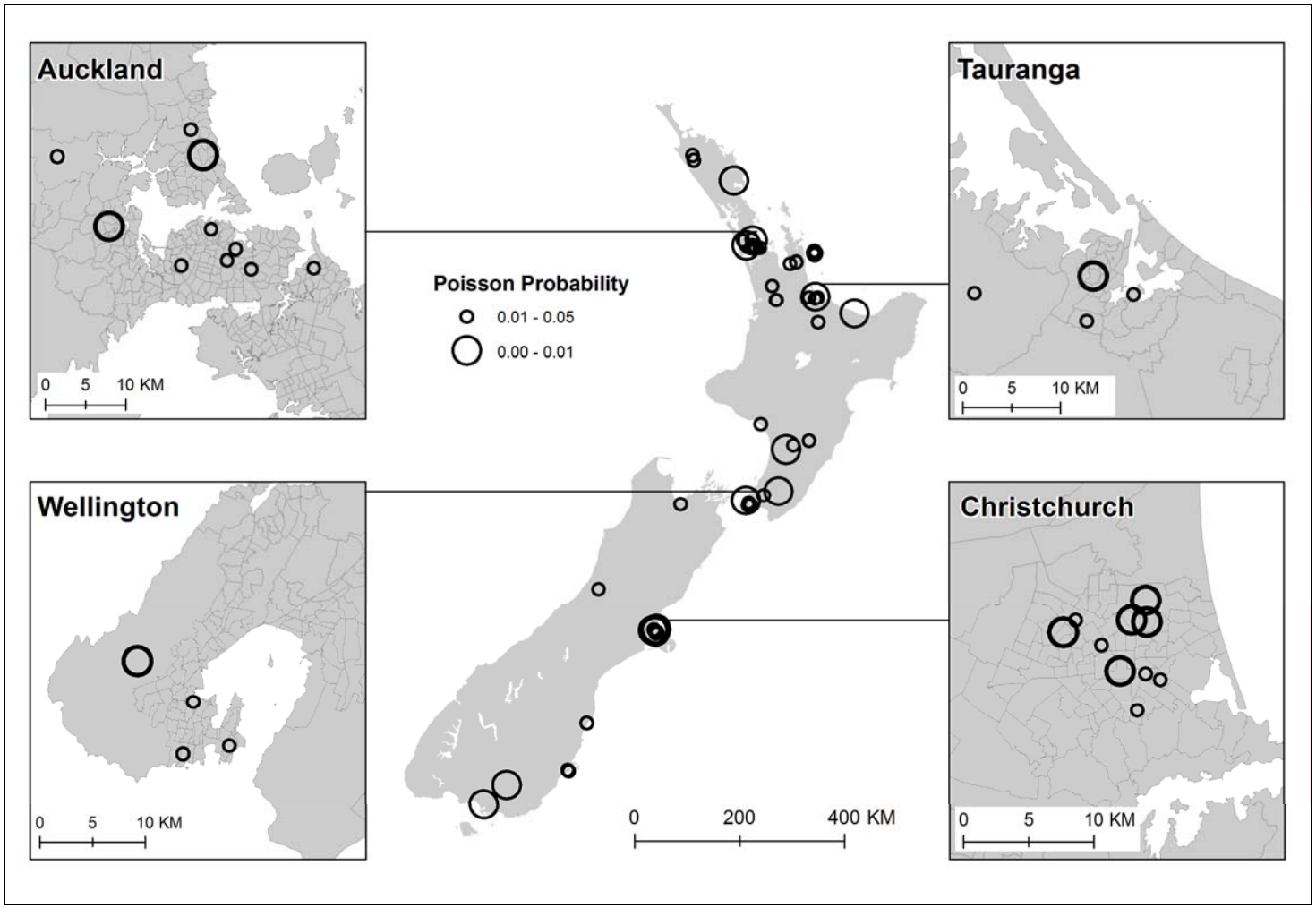

Figure 7.13: Poisson probabilities (less than 0.05) of ALL 1980-2004

7.3.3 Spatial-temporal clustering of ALL

Cluster analysis is a useful analytical technique because it considers the spatial dependence of ALL cases rather than treating each area as separate and unrelated to neighbouring areas. Furthermore, cluster techniques can effectively add a temporal element to spatial analyses. In this research, spatial-temporal cluster analyses adopting the Poisson probability model were conducted using the spatial scan statistic developed by Kulldorff (1997, 2002, 2006).

At the CAU-level, separate runs were conducted for all cases, cases disaggregated by sex, and cases disaggregated by age group at diagnosis. Cases were then aggregated to the smaller meshblock (MB)-level with the aim of detecting clusters that might occur over smaller spatial and temporal scales. At this spatial-level, separate runs were carried out for all cases, and cases disaggregated by age group at diagnosis. Table 7.8 below summarises the results of the CAU-level analyses. Significant clustering of ALL cases was observed for all cases and when males were analysed separately.

Table 7.8: Summary of ALL clusters analyses results at the CAU-level

Analysis	**P-value of most likely cluster**
All cases:	
0-14 years, males & females	<0.05
Cases disaggregated by sex:	
0-14 years, males only	<0.05
0-14 years, females only	<1.00
Cases disaggregated by age group:	
0-4 years, males & females	<0.50
5-9 years, males & females	<0.10
10-14 years, males & females	<0.50

7.3.3.1 All cases

Analysing all of the 781 cases together identified a significant ($p < 0.05$) cluster in Northern Christchurch (Table 7.9 and Figure 7.14). The cluster covered an area with a radius of 1.89km and included the CAUs of Marshland, Shirley East and Shirley West. In this area a total of 10 new cases of childhood ALL were diagnosed between 1992 and 2001 where less than one new case was expected (relative risk =10.48). A number of secondary clusters were also revealed in this analysis, such as the cluster in central Christchurch (Figure 7.14). However, the p-values for all of these clusters were greater than 0.1 (Table 7.9).

Table 7.9: Details of clusters of children with ALL aged 0-14 years at diagnosis, 1980-2004: CAU-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Observed	Expected	Relative Risk
1	Northern Christchurch	1.89	1992/1/1	2001/12/31	0.0392	10	0.96	10.48
2	Wellington	5.77	1986/1/1	1997/12/31	0.1185	26	7.87	3.38
3	Tauranga	1.67	1992/1/1	1994/12/31	0.2606	6	0.31	19.53
4	Central Christchurch	1.85	1982/1/1	1993/12/31	0.2633	12	1.90	6.41
5	Te Awamutu	9.06	1983/1/1	1983/12/31	0.7092	6	0.42	14.50
6	Cape Rodney & Waipu	22.08	1997/1/1	2001/12/31	0.9627	6	0.52	11.72
7	Invercargill	0.00	1982/1/1	1983/12/31	0.9637	2	0.01	276.69
8	Auckland	11.77	2000/1/1	2001/12/31	0.9998	23	8.85	2.65

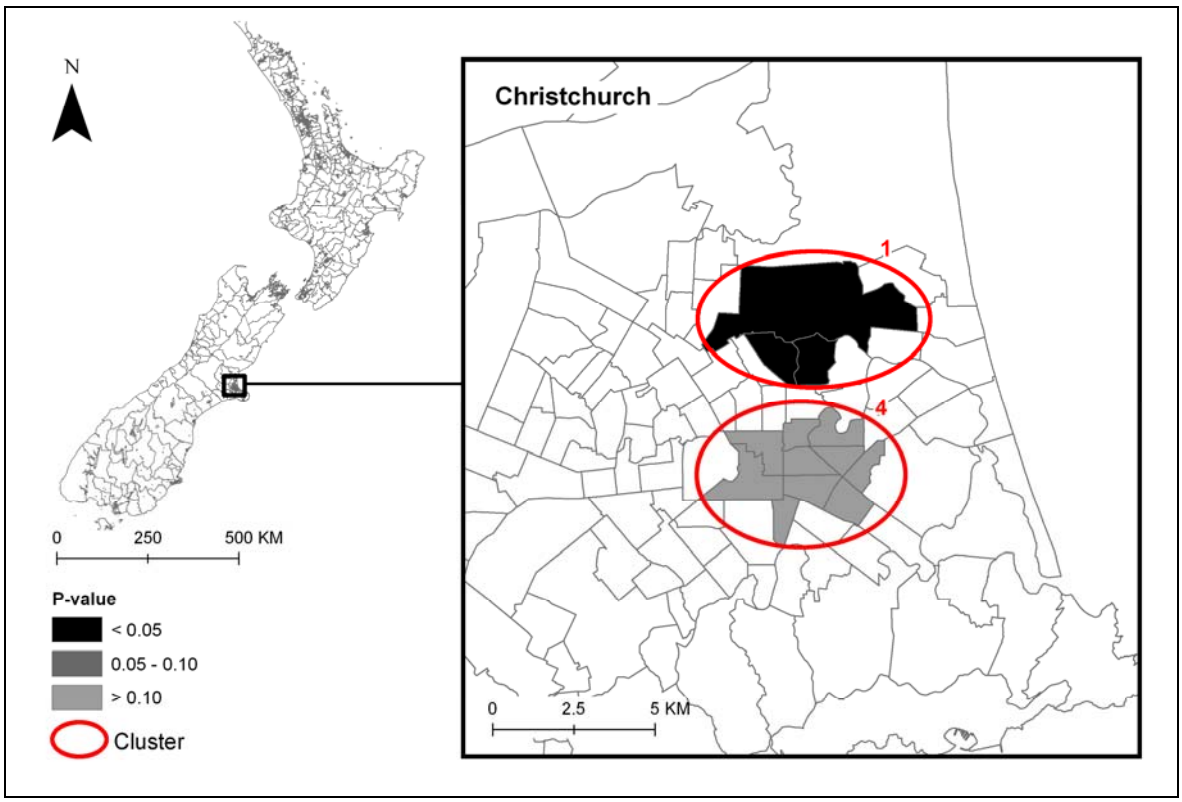

Figure 7.14: Clusters of children with ALL aged 0-14 years at diagnosis for the period 1980-2004: CAU-level

7.3.3.2 Clusters by sex

There was no evidence of spatial-temporal clustering of female ALL cases when analysed separately (Table 7.8). However, significant spatial-temporal clustering of male ALL cases was noted (Table 7.10 and Figure 7.15). The most likely cluster identified occurred in Tauranga during the period 1992-1994 (inclusive). The cluster area included five CAUs located to the south west of central Tauranga with a radius of 1.67km. During the three year period, a total of six new ALL cases developed in males aged 0-14 years at diagnosis. According to the age and sex structure of the population at risk, 0.19 cases were expected during this time (relative risk = 32.61). This cluster was statistically significant at the 95 percent confidence level. Eleven secondary clusters were also observed, with p-values ranging from 0.056 in central Christchurch to 0.999 in the Thames Coromandel. The central Christchurch cluster occurred over a twelve year period and included a total of 13 new male cases of ALL where only 1.93 cases were expected. This cluster included the CAUs of Shirley East and Shirley West which were part of the significant cluster of male and female cases between 1992 and 2001 (Figure 7.15).

Table 7.10: Details of clusters of male children with ALL aged 0-14 years at diagnosis, 1980-2004: CAU-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Observed	Expected	Relative Risk
1	Tauranga	1.67	1992/1/1	1994/12/31	0.0215	6	0.19	32.61
2	Central Christchurch	2.51	1985/1/1	1996/12/31	0.0566	13	1.93	6.92
3	Hamilton	2.14	1983/1/1	1984/12/31	0.6277	5	0.23	21.51
4	Otakaro Park Invercargill	0.00	1982/1/1	1983/12/31	0.7135	2	0	463.04
5	Wellington	1.94	1986/1/1	1987/12/31	0.7802	4	0.12	32.66
6	North Shore	2.17	1999/1/1	2001/12/31	0.9804	5	0.34	14.88
7	SW Auckland	1.23	1987/1/1	1990/12/31	0.9825	5	0.34	14.77
8	Mangawhai Kaipara/Rodney	22.08	1999/1/1	2001/12/31	0.9908	4	0.18	22.31
9	Swanson Auckland	0.00	1986/1/1	1986/12/31	0.9974	2	0.01	170.27
10	SE Auckland	4.22	2000/1/1	2000/12/31	0.9990	5	0.4	12.79
11	Thames Coromandel	25.39	1981/1/1	1986/12/31	0.9999	4	0.22	18.35

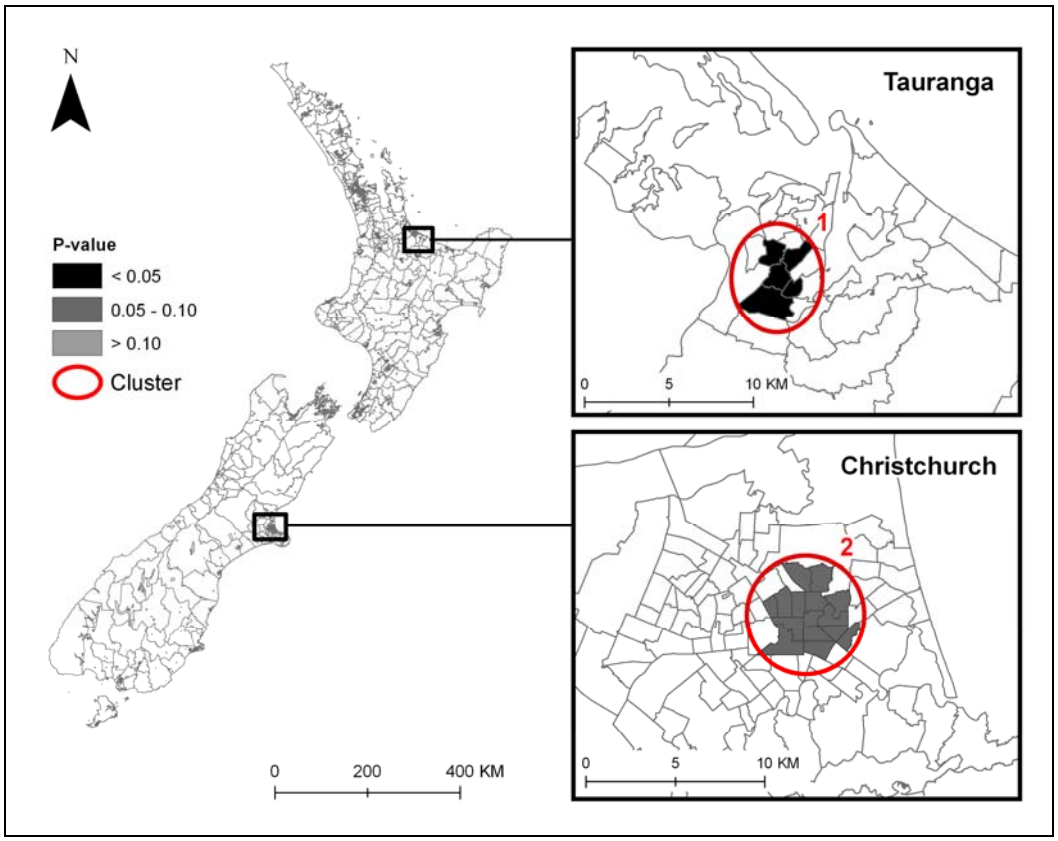

Figure 7.15: Clusters of male children with ALL aged 0-14 years at diagnosis for the period 1980-2004: CAU-level

7.3.3.3 Clusters by age group

Disaggregating the ALL cases by age group at diagnosis did not reveal any significant spatial-temporal clusters (Table 7.8). In the 5-9 year analysis, a total of seven possible clusters were identified (Table 7.11) and the most likely cluster (number 1) occurred in the CAU of Te Reti, a suburb to the south west of central Tauranga (Figure 7.16). This suburb was identified as being part of a significant cluster in the male only analyses (Figure 7.15) for the same time period (1992-1994). In the age group analysis, a total of three new ALL cases developed in male and female children aged 5-9 years at diagnosis, where 0.01 cases were expected. The relative risk of developing ALL in the area during the two year period was thus very high (253). The p-value for this cluster was 0.0536 and was therefore statistically significant at the 90 percent confidence level.

Table 7.11: Details of clusters of children with ALL aged 5-9 years at diagnosis, 1980-2004: CAU-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Observed	Expected	Relative Risk
1	Te Reti - Tauranga	0.00	1992/1/1	1994/12/31	0.0536	3	0.01	253.05
2	Selwyn & Waimakariri DC	53.05	1988/1/1	1990/12/31	0.9153	3	0.05	55.61
3	Kilbirnie - Wellington	0.95	1991/1/1	1996/12/31	0.9368	3	0.06	53.15
4	Central Christchurch	4.72	1999/1/1	2002/12/31	0.9595	7	0.84	8.58
5	Western Wellington	1.03	1997/1/1	1997/12/31	0.9976	2	0.01	142.96
6	Oamaru	0.00	1991/1/1	1993/12/31	0.9992	2	0.02	132.10
7	Ruapehu	21.21	2002/1/1	2002/12/31	0.9999	2	0.02	101.83

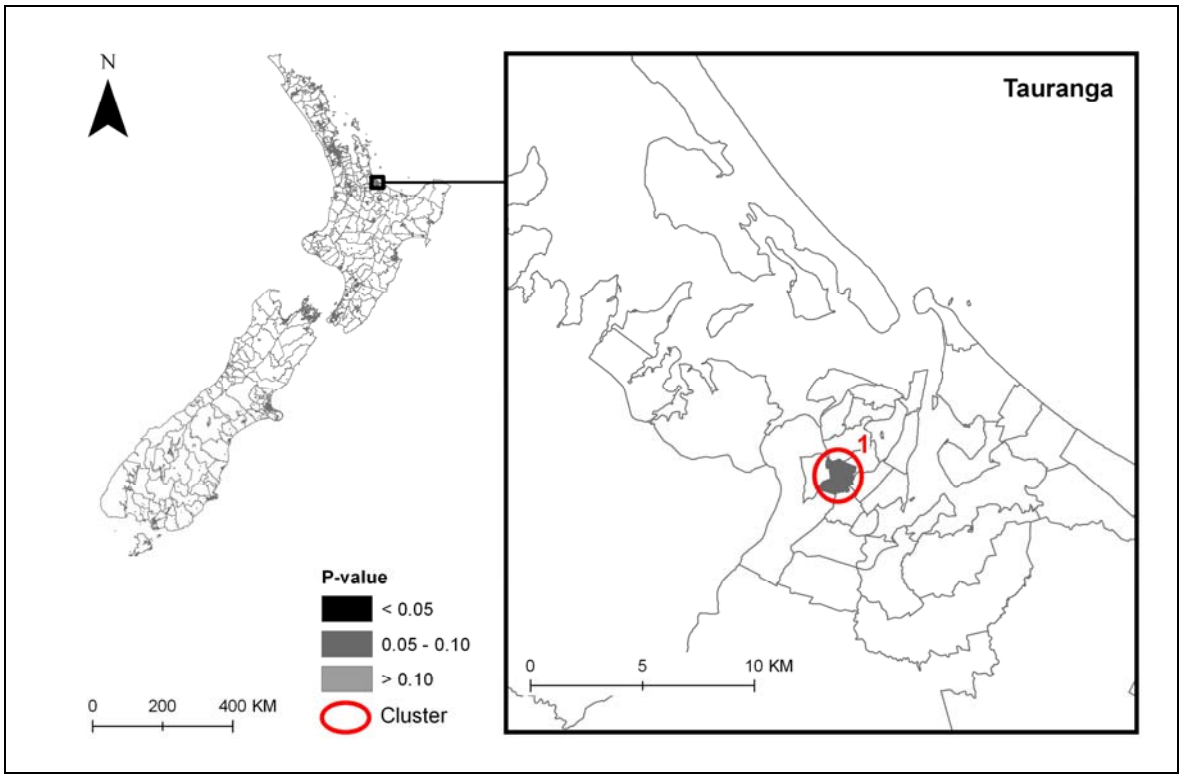

Figure 7.16: Clusters of children with ALL aged 5-9 years at diagnosis for the period 1980-2004: CAU-level

7.3.3.4 Meshblock-level analyses

Due to the large number of MBs in New Zealand, separate MB-level analyses were carried out for the North and South Island (Table 7.12). No significant clusters were observed in the South Island during the entire study period in any of the analyses. Conversely, at least one significant cluster was identified in each of the North Island runs. For example, when all of the cases were

analysed together, two significant clusters were revealed (Table 7.13 and Figure 7.17). The most likely cluster (number 1) occurred in the Pinehill area of North Shore in 2003, where two new cases of ALL were observed when no new cases were expected. Due to the small number of observed and expected cases, the relative risk for this cluster was extremely high (17,667). However its p-value was very low; this cluster was statistically significant at the 99 percent confidence level. A similarly significant cluster was also noted in central Hamilton in 1983 (Table 7.13 and Figure 7.17). Again, two new cases of ALL were observed compared to no new cases expected.

Table 7.14 gives the details of the most likely clusters found by the MB age group analyses. Central Hamilton was again highlighted as a significant cluster, this time of children aged 0-4 years at ALL diagnosis. In 1983, two children in this age group developed ALL where no new cases were expected, again resulting in an extremely high relative risk value (33,385). As with the all-age cluster in Hamilton, this cluster of 0-4 year olds was significant at the 99 percent confidence level. For children aged 5-9 years at diagnosis, a significant cluster ($p< 0.05$) was observed in the suburb of Te Reti in Tauranga in 1994, based on four observed and no expected cases. Pinehill in North Shore noted a highly significant cluster ($p<0.01$) of ALL cases aged 10-14 years at diagnosis in 2003. Two new cases of ALL occurred when none were expected according to the population structure of the area. The relative risk for this cluster was 98,951. No significant secondary clusters were observed in the age group analyses.

Table 7.12: Summary of ALL cluster analyses results at the MB-level

Analysis	**P-value of most likely cluster**
All cases:	
0-14 years, North Island	<0.01
0-14 years, South Island	<0.50
Cases disaggregated by age group:	
0-4 years, North Island	<0.01
0-4 years, South Island	<0.50
5-9 years, North Island	<0.05
5-9 years, South Island	<0.10
10-14 years, North Island	<0.01
10-14 years, South Island	<0.50

Table 7.13: Details of clusters of children with ALL aged 0-14 years at diagnosis, 1980-2004: MB-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Observed	Expected	Relative Risk
1	Pinehill-North Shore	0.00	2003/1/1	2003/12/31	0.0035	2	0.00	17,667.59
2	Central Hamilton	0.00	1983/1/1	1983/12/31	0.0035	2	0.00	17,667.59
3	SE Tauranga	1.87	1994/1/1	1994/12/31	0.1018	4	0.05	86.68
4	Forest Hill - North Shore	0.22	2001/1/1	2001/12/31	0.4755	2	0.00	1,104.22
5	Swanson - Waitakere City	0.00	1986/1/1	1986/12/31	0.7812	2	0.00	679.52
6	Marton - Whanganui	0.46	1982/1/1	1982/12/31	0.8804	2	0.00	552.11
7	Cape Rodney	2.51	2000/1/1	2000/12/31	0.9348	2	0.00	471.94
8	Wellington City	2.02	1987/1/1	1987/12/31	0.9891	4	0.14	29.47
9	Strathmore Park - Wellington	0.33	1997/1/1	1997/12/31	0.9956	2	0.01	294.46
10	Omokoroa - Western Bay of Plenty	2.09	2004/1/1	2004/12/31	0.9976	2	0.01	269.68
11	SE Auckland	4.75	2000/1/1	2000/12/31	0.9998	7	0.84	8.41
12	SW Rotorua	1.23	2003/1/1	2003/12/31	0.9998	3	0.06	47.91

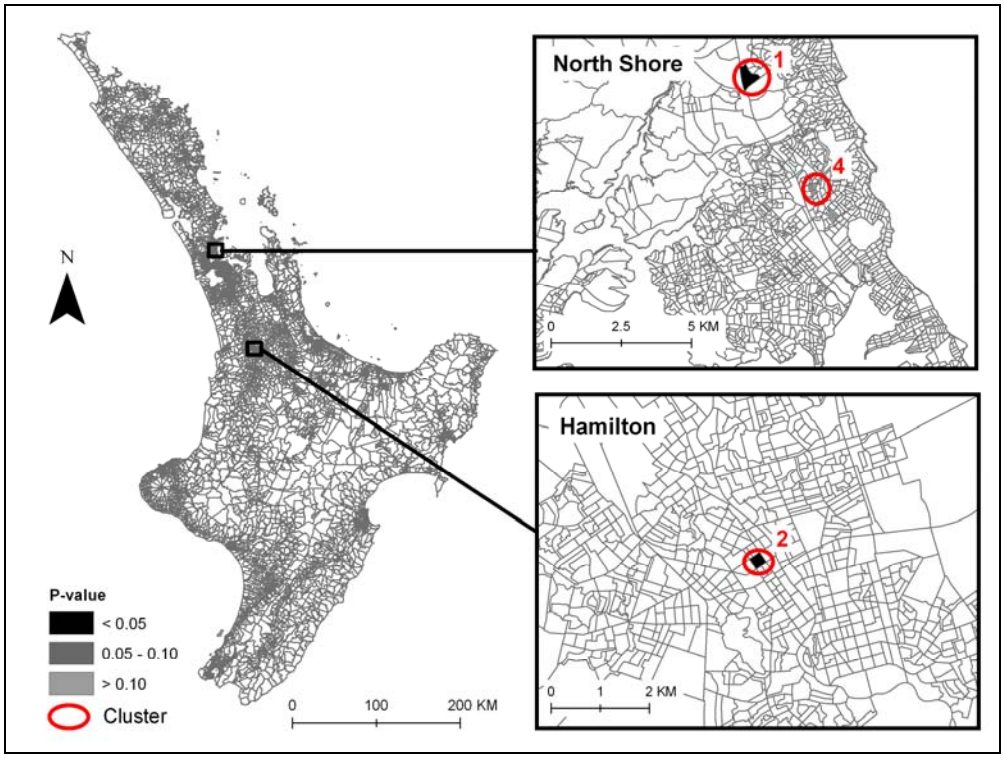

Figure 7.17: Clusters of children with ALL aged 0-14 years at diagnosis for the period 1980-2004: MB-level, North Island

Table 7.14: Most likely ALL clusters by age group at diagnosis 1980-2004: MB-level, North Island

Age group (years)	Cluster Location	Radius (km)	Start Date	End Date	P-value	Observed	Expected	Relative Risk
0-4	Central Hamilton	0.00	1983/1/1	1983/12/31	0.0013	2	0.00	33,385.10
5-9	Te Reti - Tauranga	0.00	1994/1/1	1994/12/31	0.0479	2	0.00	3719.62
10-14	Pinehill North Shore	0.00	2003/1/1	2003/12/31	0.0001	2	0.00	98,951.08

The descriptive and geographical analyses reported so far have revealed a number of interesting patterns in the occurrence of ALL in New Zealand children. In summary, ALL has been found to be higher in younger children (0-4 years) compared to older children (10-14 years), in males compared to females, and in children of European ethnicity compared to children in the other main ethnic groups in New Zealand. Incidence of ALL increased slightly between 1980 and 2004 and varied substantially between areas. Significantly lower ALL incidence was found in the most deprived tenth and fifth of all CAUs in New Zealand, and lower incidence was noted in the most rural areas of the country. A number of geographical analyses confirmed this finding, showing an excess of ALL cases in the main urban centres of both islands. Furthermore, significant spatial-temporal clustering of cases was observed in urban areas. Northern Christchurch, central Tauranga, North Shore and central Hamilton were identified as areas having significant clusters of this disease.

7.4 ALL and population mixing

The findings described in the first half of this chapter have addressed the first main aim of the thesis: to determine the geographical epidemiology of childhood ALL in New Zealand for the period 1980 to 2004. The remainder of the chapter addresses the second aim of the thesis and thus concentrates on examining whether there is a relationship between population mixing and childhood ALL in the New Zealand setting. Can the geographical and temporal patterns observed in ALL be explained by variations in population mixing levels around the country and over time?

7.4.1 Exploratory analysis

7.4.1.1 ALL by population change and migration change quintile

A number of exploratory data analyses were carried out to examine the relationship between childhood ALL and population mixing. CAUs were ranked and grouped into quintiles of relative population change between 1981 and 2001 (Table 5.10, chapter 5). For each population change quintile, the observed and expected cases of ALL were summed, and the corresponding SIRs, upper and lower confidence intervals and chi-square values were calculated (Table 7.15). Quintiles 2-4 had SIRs greater than 100 suggesting higher than average ALL cases in areas which increased in population by up to 46.73 percent or declined by up to 7.97 percent. Quintiles 1 (highest population growth) and 5 (most population decline) both had SIRs below 100 and thus witnessed fewer than expected cases over the study period. It should be noted that only the SIR for quintile 5 was statistically significant (83.30, CI = 69.41-97.20) with a chi-square value of 3.85. Thus, areas which experienced considerable population losses between 1981 and 2001 observed significantly fewer ALL cases than expected.

Table 7.15: ALL SIRs, chi-square values and CIs by quintile of population change, 1980-2004

Quintile of Population Change	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
1 High growth	138	144.05	95.80	80.15	111.44	0.25	NO
2	163	149.21	109.24	91.72	126.77	1.28	NO
3	179	174.60	102.52	87.31	117.73	0.11	NO
4	186	174.04	106.87	91.00	122.75	0.82	NO
5 High decline	115	138.05	83.30	69.41	97.20	3.85	YES

The analysis was repeated using quintiles of change in the proportion of migrants in each area in 2001 compared with 1981 (Table 5.11, chapter 5). The ALL SIRs generally decreased over the quintiles of migration change, although the pattern was not linear (Figure 7.18). The highest SIRs were recorded in quintiles 1 and 2 (109.36 and 102.69 respectively), implying that an excess of ALL cases were observed in areas whose percentage of incoming migrants increased by between 15.78 and 355.82 percent. Fewer than expected cases were observed in quintiles 3 and 5: areas which either increased or decreased slightly (quintile 3), or which witnessed considerable decreases (quintile 5) in in-migration. The confidence intervals for the SIRs of each migration change quintile included 100 (Figure 7.18).

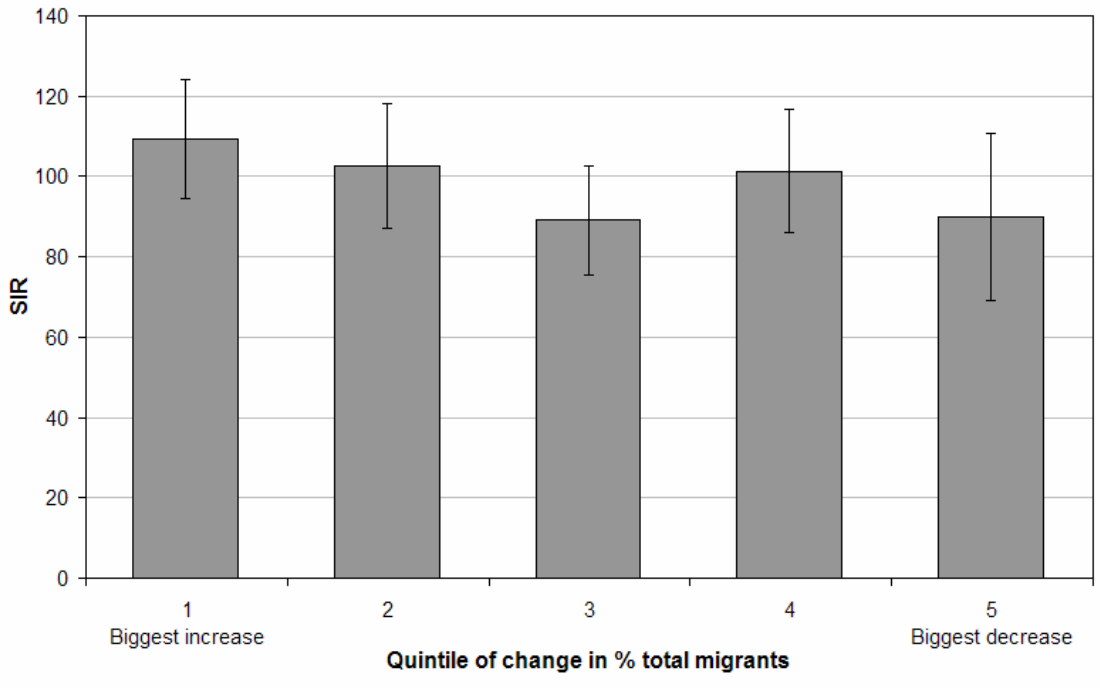

Figure 7.18: ALL SIRs by quintile of change in the percentage of total migrants, 1980-2004

7.4.1.2 Correlation analysis

Correlation analyses were also carried out at the CAU-level to explore the relationships between ALL and population mixing. The Spearman's rank order correlation coefficients were calculated to summarise the bivariate associations between ALL and each of the population mixing measures and control variables.

The majority of the population mixing measures were weakly correlated with the ALL count for the period 1980-2004 (Table 7.16). The strongest association occurred between ALL and the change in the number of total migrants. This variable had a correlation coefficient of 0.208, suggesting a weak positive relationship with childhood ALL ($p<0.01$). As the number of migrants in an area increased, so too did the number of ALL cases. A slightly weaker positive correlation was observed between ALL and the change in the percentage of total ($r = 0.190$) and child ($r = 0.194$) migrants' variables. The change in the number of child migrants was also weakly associated with ALL count during the study period. The only other significant correlation was shown for the change in the percentage of overseas migrants. However, the correlation coefficient was very close to zero ($r = 0.074$). Population change, change in migrant diversity and change in the percentage of overseas visitors, were not significantly associated with ALL count in the correlation analyses.

Table 7.16: Spearman's rank order correlation coefficients between ALL count and population mixing variables, 1980-2004

Population mixing variable	Correlation Coefficient
Change in the number of total migrants	0.208**
Change in the number of child migrants	0.167**
Population change	0.023
Change in the percentage of total migrants	0.190**
Change in the percentage of child migrants	0.194**
Change in the percentage of overseas migrants	0.074**
Change in the percentage of overseas visitors	-0.039
Change in migrant diversity	-0.011

** denotes statistically significant at the 99% confidence level
* denotes statistically significant at the 95% confidence level

As well as providing useful information regarding the strength and direction of relationships between two variables, correlation analysis is also an important precursor to regression modelling. It allows the examination of the relationship between variables used in regression models and can thus be employed to identify any potential confounding occurring within the models. Spearman's rank order correlation coefficients between ALL and a number of control variables were thus also examined (Table 7.17). Additionally, the associations between the population mixing variables and these control variables were inspected (Table 7.18).

The count of ALL cases between 1980 and 2004 was significantly and positively ($r = 0.414$) correlated with the population of 0-14 year olds at the mid-point of the study period (1991). A similar but slightly weaker association was noted for population density in 1991 ($r = 0.235$). Very small positive correlation coefficients were found between ALL and both the deprivation score and the percentage of households with more than six usual members. A weak negative association was noted for the percentage of Europeans' variable, and no significant relationship was found between ALL and the percentage of households with less than three usual members (Table 7.17).

Table 7.17: Spearman's rank order correlation coefficients between ALL count and control variables, 1980-2004

Control variable	Correlation Coefficient
Population 1991	0.414**
% European	-0.113**
Population Density	0.235**
Deprivation Score	0.070**
% Households < 3 people	-0.006
% Households > 6 people	0.052*

** denotes statistically significant at the 99% confidence level
* denotes statistically significant at the 95% confidence level

A number of significant correlations were noted between the population mixing measures and the control variables (Table 7.18). At the CAU-level, the population count of 0-14 year olds in 1991 was positively correlated with the change in the number of total and child migrants, and the change in the percentage of total, child and overseas migrants. Population count had a weak negative association with the change in the percentage of overseas visitors' variable. A similar pattern of association was noted for the population density variable, and the percentage of households with less than three usual members. The percentage of European residents in each CAU had a weak inverse association with the change in the number/percentage of child/total migrants' variables, and a weak positive association with population change, change in the percentage of overseas visitors and migrant diversity. The exact opposite relationships were noted for deprivation score. The percentage of households with more than six usual members' variable was negatively associated with every population mixing measure.

Table 7.18: Spearman's rank order correlation coefficients between the control and population mixing variables, 1980-2004

	Number of total migrants	Number of child migrants	Population change	Change % migrants	Change % child migrants	Change % overseas migrants	Change % overseas visitors	Change migrant diversity
Population 1991	0.378**	0.274**	-0.009	0.375**	0.390**	0.175**	-0.097**	-0.020
% European	-0.084**	-0.070**	0.081**	-0.181**	-0.169**	-0.010	0.169**	0.113**
Population Density	0.314**	0.241**	-0.045	0.383**	0.369**	0.267**	-0.131**	0.031
Deprivation Score	0.068**	0.092**	-0.239**	0.283**	0.275**	0.005	-0.112**	-0.136**
% Households < 3 people	0.183**	0.174**	-0.023	0.254**	0.225**	0.096**	0.061*	-0.003
% Households > 6 people	-0.126**	-0.121**	-0.080**	-0.058*	-0.048*	-0.075**	-0.097**	-0.098**

** denotes statistically significant at the 99% confidence level
* denotes statistically significant at the 95% confidence level

7.4.2 Regression modelling

Examining simple correlations and SIRs by quintiles of population or migration change are helpful to begin exploring the relationships between ALL and population mixing. However, only one variable is considered at a time in these analyses, and similar to many other cancers, ALL is likely to have multiple influencing factors (Greaves, 2006). Consequently, multivariate Poisson regression analyses were employed to ascertain the relationships between ALL and population mixing at the CAU-level, whilst controlling for other potential predictors of the disease.

The modelling was undertaken in six stages (Figure 7.19). The results of steps 1 and 2 were discussed in chapter 5 (Data and methods), and the results of steps 3 through to 6 are reported in the remainder of this chapter. Univariate model results are presented first followed by a description of how the multivariate models were formulated. The final section details the results of the multivariate models. Analyses were conducted for the study period as a whole (1980-2004), two time periods of 12 and 13 years (1980-1991 and 1992-2004), and four time periods of 6/7 years (1980-1986, 1987-1992, 1993-1998 and 1999-2004) (Table 7.19).

Summary of the modelling strategy:

1. Initially, the appropriate type of probability distribution was chosen to model the data. Examination of the descriptive statistics revealed that the Poisson distribution was an appropriate starting model. Thus, at the outset, Poisson models were fitted to random combinations of variables. These models were then statistically compared with negative binomial and zero-inflated models to determine which model better fitted the data.
2. After the appropriate type of regression model was selected, each explanatory variable was examined to check whether they should be transformed for inclusion in the models.
3. In the first instance, models containing only an intercept were fitted in order to provide a measure of the variation in the dependent variable around its mean.
4. Next all of the explanatory variables were then modelled separately. The age and sex-specific population at the mid-point of the study period was included as an exposure variable in each model and robust standard errors were calculated to account for within area clustering.
5. Using theory guided by the literature and the results of the univariate control models for the 12/13 and 25 year analyses, a base model and alternative multivariate models were constructed (Figure 7.20). Where univariate results for the year group being analysed indicated additional significant control variables, the base and alternative models were altered accordingly.
6. Each population mixing measure was then added separately to each of these models.

Figure 7.19: Summary of regression modelling strategy
(See chapter 5 for more detail)

Table 7.19: ALL analysis time periods

Time period	Number of years	Total cases
1980-2004	25	781
1980-1991	12	342
1992-2004	13	439
1980-1986	7	183
1987-1992	6	185
1993-1998	6	194
1999-2004	6	219

7.4.2.1 Univariate model results

This section details the results of the univariate regression analyses which examined how ALL varied by time period, various individual- and area-level control variables, and a number of area-level population mixing measures. First, tests for over-dispersion and zero-inflation were conducted and the appropriate models for each analysis were decided upon.

Data for the change over time analyses and both the 1980-2004 and 1992-2004 analyses were significantly over-dispersed. As a result, negative binomial regression models were employed to examine these data. The 1980-1991 dataset showed no significant over-dispersion so Poisson regression models could be utilised. Using zero-inflated models did not significantly improve the model fit for any of these analyses and were therefore not applied at this stage. Using the appropriate model type, each variable was modelled alone with the age- and sex-specific population at risk incorporated as an exposure variable. Associations between the ALL counts and the explanatory variables were calculated as incidence rate ratios (IRRs). IRRs provide a comparison of disease rates between different populations/areas and are obtained by exponentiating the Poisson regression coefficient. For continuous variables, IRRs greater than one suggest an increased relative risk of ALL incidence in populations/areas where values of the independent variable are high. IRRs below one reveal a lower relative risk in populations/areas where values of the independent variable are high. For categorical variables, IRRs greater than one show an increased relative risk of ALL for the category in question when compared to the base category.

Change over time

Change in ALL incidence over time was examined first. As noted in section 7.2 of this chapter, the incidence of ALL increased slightly during the study period (1980-2004). Negative binomial regression models were used to test whether this increase was statistically significant. ALL cases were aggregated by year group at diagnosis for the periods: 1980-1991 and 1992-2004 (12/13 years); and 1980-1986, 1987-1992, 1993-1998, and 1999-2004 (6/7 years) (Table 7.19).

The ALL IRRs produced by the negative binomial regression models increased over time (Tables 7.20 and 7.21). Furthermore, a significant increase in ALL was observed in the 12/13 year analyses (Table 7.20). Around 22 percent more cases were observed between 1992 and 2004 when compared to the period 1980-1991 (IRR = 1.215, CI = 1.051-1.404). Disaggregating

the cases into 6/7 year groups at diagnosis revealed a similar temporal trend in incidence (Table 7.21). Again, the highest incidence was observed in the most recent period (1999-2004) and the lowest incidence in the earliest time period (1980-1986). With an IRR of 1.149, approximately 15 percent more cases of ALL were diagnosed between 1999 and 2004 compared to the period 1980-1986. The middle two periods (1987-1992 and 1993-1998) had similar IRRs, suggesting a levelling out in incidence during this time, which was also noted in section 7.2. However, the confidence intervals for all of the IRRs included one, and thus the trend over the 6/7 year time periods was not statistically significant.

Table 7.20: Results of the ALL univariate negative binomial regression analyses for year group at diagnosis (12/13 years)

Year Group	IRR	Robust SE	P-value	LCI	UCI
1980-1991 (Base Category)	1.000				
1992-2004	1.215	0.090	0.008	1.051	1.404

Table 7.21: Results of the ALL univariate negative binomial regression analyses for year group at diagnosis (6/7 years)

Year Group	IRR	Robust SE	P-value	LCI	UCI
1980-1986 (Base Category)	1.000				
1987-1992	1.058	0.111	0.593	0.861	1.301
1993-1998	1.048	0.109	0.653	0.854	1.285
1999-2004	1.149	0.116	0.170	0.942	1.401

Control variables

After considering how ALL varied by year group of diagnosis, differences by a number of control variables were considered. Results are presented for 12/13 and 25 year time periods (Table 7.22) as these were used to inform the 6/7 year multivariate analyses. First, models containing only an intercept were fitted in order to determine the mean number of cases diagnosed for each study period. During the whole study period, an average of 0.076 new cases of ALL were observed in each CAU across New Zealand (CI = 0.070-0.082). Dividing the study period roughly in half revealed a higher average number of cases (0.043) for the period 1992-2004 compared to 1980-1991 (0.033 cases). This result concurs with the findings of the 12/13 year group analyses (Table 7.20).

Table 7.22: Univariate results of the ALL regression models by 12/13 & 25 year group and control variable

Control variable	1980-1991 Poisson	1992-2004 Negative Binomial	1980-2004 Negative Binomial
None (null model)	0.033* (0.030-0.037)	0.043* (0.039-0.047)	0.076* (0.070-0.082)
Age group category			
0-4 years (base)	1.000	1.000	1.000
5-9 years	0.411* (0.318-0.531)	0.503* (0.403-0.627)	0.525* (0.444 - 0.622)
10-14 years	0.260* (0.195-0.347)	0.329* (0.254-0.426)	0.344* (0.283 - 0.418)
Sex category			
Male (base)	1.000	1.000	1.000
Female	0.769* (0.621-0.953)	0.733* (0.604-0.891)	0.745* (0.644 - 0.863)
% European	1.005 (0.998-1.011)	1.003 (0.998-1.008)	1.004* (1.000 - 1.008)
Deprivation	1.000 (0.999-1.001)	0.999 (0.998-1.000)	0.999 (0.998 - 1.000)
Population density	1.000* (1.000-1.000)	1.000 (0.999-1.000)	1.000* (1.000 - 1.000)
% Households <3 people	1.023* (1.011-1.035)	1.005 (0.995-1.015)	1.014* (1.006 - 1.022)
% Households >6 people	0.962* (0.936-0.990)	0.989 (0.971-1.008)	0.974* (0.957 - 0.991)
Urban/rural category (7)			
1-Main urban areas (base)	1.000	1.000	1.000
2-Satellite urban areas	0.897 (0.476-1.689)	1.183 (0.715-1.958)	1.044 (0.699 - 1.559)
3-Independent urban comm.	1.016 (0.751-1.374)	0.801 (0.580-1.105)	0.865 (0.691 - 1.083)
4-Rural high urban influence	0.758 (0.358-1.608)	1.305 (0.802-2.125)	1.106 (0.733 - 1.668)
5-Rural mod urban influence	0.559 (0.264-1.185)	1.080 (0.657-1.775)	0.842 (0.555 - 1.276)
6-Rural low urban influence	0.724 (0.476-1.101)	0.779 (0.526-1.154)	0.725* (0.542 - 0.969)
7-Highly remote/rural	0.592 (0.244-1.437)	0.904 (0.423-1.933)	0.713 (0.398 - 1.274)
Urban/rural category (2)			
1-Predominantly urban (base)	1.000	1.000	1.000
2-Predominantly rural	0.680* (0.486-0.950)	0.978 (0.754-1.269)	0.826 (0.674 - 1.012)

IRRs are reported with 95% confidence intervals in the parentheses.
* denotes a p-value of <0.05.

In both the 12/13 and 25 year analyses, a strong inverse relationship between age group at diagnosis and ALL was observed (Table 7.22). For all of the time periods, children in the youngest age group of 0-4 years at diagnosis (the base category) witnessed significantly more ALL cases than those in the two older age groups. Furthermore, the IRRs revealed a dose-response relationship which was consistent over time. For example, compared to children aged 0-4 years, 47.5 percent fewer cases of ALL were diagnosed in 5-9 year olds and 65.6 percent fewer cases were diagnosed in 10-14 year olds, for the period 1980-2004. The greatest difference between the age groups occurred in the earliest period (1980-1991), when 74.0 percent fewer ALL cases were observed in the oldest age group (10-14 years) compared to the youngest age group (0-4 years). None of the confidence intervals for the IRRs included one. ALL was also consistently higher in male children. Female ALL IRRs were significantly lower than male IRRs

in all of the three time periods. The largest difference between the two sexes occurred between 1992 and 2004 when around 26.7 percent fewer female, compared to male, ALL cases were observed.

Interesting relationships were also found between ALL and a number of the area-level control variables. ALL incidence was higher in areas with a high percentage of European residents for all of the temporal periods. Between 1980 and 2004 the IRR for this variable (1.004) was statistically significant at the 95 percent confidence level (CI = 1.000-1.008). The most deprived areas of New Zealand had lower incidence of ALL for the periods 1992-2004 and 1980-2004. For the study period as a whole, the IRR for this variable was 0.999 (CI = 0.998-1.000) and was significant at the 90 percent confidence level. Childhood ALL was significantly higher in the most densely populated CAUs of the country during the first half of the study (1980-1991) and the whole 25 year period (1980-2004). The IRR of 1.0001 (4dp) suggests that as the number of people living per square kilometre increased, the incidence of ALL also increased slightly. The household overcrowding variables were both significantly associated with ALL incidence during the study period, especially in the first 12 years. Areas with a high proportion of households with less than three usual members had significantly higher ALL rate ratios. Compatible with this result is the finding that areas with a high proportion of houses with more than six usual members (overcrowded areas), had significantly lower incidence of this disease.

There were no notable trends in ALL IRRs across the seven urban/rural categories for any of the time periods. However, the IRRs for the most rural areas (category 6 and 7 areas) were consistently below one, suggesting that the incidence of ALL was lower in the most rural parts of New Zealand compared to the main urban areas (the base category). Between 1980 and 2004, the IRR (0.725, CI = 0.542-0.969) for rural areas with a low urban influence was statistically significant and suggests that 27.5 percent fewer cases were noted in these areas compared to the main urban areas. Aggregating areas into two urban/rural categories, showed consistently lower IRRs in the predominantly rural areas compared to the predominantly urban areas. However, the difference between the two categories was only significant between 1980 and 1991.

The associations found between ALL and the control variables were generally consistent across the time periods. ALL was highest in children aged 0-4 years at diagnosis, in males compared to females, in areas with a high percentage of Europeans, in the least deprived areas, in the most densely populated areas, and in areas with a high proportion of small households (less than three usual members). In addition, ALL was significantly lower in areas where there was more

household overcrowding (more than six usual members), and in rural areas compared to urban areas.

Population mixing change variables

In the univariate 12/13 and 25 year analyses, the majority of the population mixing change variables were not significantly associated with childhood ALL (Table 7.23). The IRRs for most of the continuous variables were very close to unity. For example, in each time period the IRR for population change was 0.999, suggesting a very weak inverse relationship with ALL. However, the 95 percent confidence intervals for this variable included one for every time period. Similarly, the change in the percentage of overseas visitors variable also had IRRs of 0.999 (CI = 0.999-1.000) for each period. The IRRs for the change in overseas migrants' variable remained just over one in every model, but were not statistically significant. The IRRs for the remaining population mixing change variables varied over time. ALL was found to be significantly higher in areas which had increased the most in the percentage of total migrants between 1980 and 1991 (IRR = 1.003, CI = 1.000-1.006). However, during the period 1992-2004 the IRR for this variable was below one, suggesting an inverse relationship with ALL for this time. For the study period as a whole, the IRR was above one. The confidence intervals for the 13 and 25 year time periods included one. A similar pattern was noted for the change in the percentage of child migrants' variable, although the IRRs did not reach statistical significance in any of the univariate models tested. The association between ALL and changes in area-level migrant diversity were close to unity in all of the three time periods, with IRRs just below one for 1980-1991 and 1992-2004, and above one for 1980-2004.

Table 7.23: Results of the ALL univariate regression analyses of the population mixing change variables for 12/13/25 yearly time periods

Population mixing change variable	1980-1991 Poisson	1992-2004 Negative Binomial	1980-2004 Negative Binomial
Population change	0.998 (0.995-1.002)	0.999 (0.998-1.001)	0.999 (0.999-1.000)
Change in % total migrants	1.003* (1.000-1.006)	0.996 (0.989-1.004)	1.001 (0.999-1.003)
Change in % child migrants	1.002 (0.999-1.005)	0.996 (0.992-1.000)	1.001 (0.999-1.002)
Change in % overseas migrants	1.001 (0.999-1.002)	1.000 (0.999-1.002)	1.000 (0.999-1.001)
Change in % overseas visitors	0.999 (0.999-1.000)	0.999 (0.999-1.000)	0.999 (0.999-1.000)
Change in migrant diversity	0.999 (0.990-1.007)	0.999 (0.991-1.008)	1.001 (0.998-1.003)
Population mixing change category:			
1. Decrease in either/both categories	1.000	1.000	1.000
2. No change in either category	1.314* (1.019-1.693)	1.271 (0.947-1.706)	1.066 (0.886-1.282)
3. Increase in diversity category	1.389 (0.939-2.056)	0.801 (0.448-1.431)	1.091 (0.843-1.411)
4. Increase in migration category	1.052 (0.770-1.437)	1.083 (0.740-1.583)	1.186 (0.979-1.436)

IRRs are reported with 95% confidence intervals in the parentheses.
* denotes a p-value of <0.05.

In general, the categorical population mixing change variable showed raised ALL incidence in category 2, 3 and 4 areas (which had not changed, or increased in population mixing category) compared to the base category areas (which had decreased in population mixing category). The only exception was an IRR of 0.801 (CI = 0.448-1.431) for category 3 areas during the period 1992-2004. Between 1980 and 1991, areas which experienced no change in population mixing category (category 2 areas) had a significantly higher ALL IRR (1.314, CI = 1.019-1.693) compared to areas which had decreased in population mixing category. For the period 1992-2004, no real trend across the population mixing categories was observable, and all of the IRRs had confidence intervals which included one. For the study period as a whole, the IRRs increased with increasing population mixing change category. The largest IRR was noted in category 4 areas. The IRR of 1.186 (CI = 0.979-1.436) implies that ALL was highest in areas which had increased the most in population mixing compared to areas which had decreased in population mixing over the study period (p-value = 0.081).

In summary, most of the population mixing change variables were not significantly related to ALL incidence in New Zealand for the periods 1980-1991, 1992-2004 and 1980-2004. Only two significant associations were noted, both in the first 12 years of the study (1980-1991). During

this time, a significantly higher incidence of ALL was observed in areas which had increased the most in total migrants, and in areas which were relatively stable (category 2 areas) in terms of both migrant diversity and in-migration.

While providing valuable insight into univariate relationships between the explanatory variables and ALL count, none of these analyses controlled for potential confounding variables. Correlation analyses previously revealed significant relationships not only between ALL and the control variables, but also between the control variables and the population mixing variables. As a result, it was necessary to examine these associations further in multivariate regression analyses.

7.4.2.2 Formulation of the multivariate models

The results of the 12/13 and 25 year univariate analyses (Tables 7.22 and 7.24) together with findings from the international literature were used to formulate a number of multivariate models (step 5 of the modelling strategy). These models aimed to assess whether the inclusion of important control variables changed the relationships between ALL and the population mixing measures, as described in the previous section. First, a simple base model was constructed, followed by a number of alternative models.

The individual variables age group at diagnosis and sex had the greatest (significant) effects on ALL (Table 7.22), and the inclusion of these variables produced the greatest reductions in deviance (Table 7.24). As a result, age group and sex formed the initial base model to which the population mixing measures were added, one at a time. In order to control for the age and sex structure of the population at risk, the age- and sex-specific population for each area was included as an exposure variable in the base model. However, the variables percentage European, deprivation score, population density, the two household overcrowding variables and urban/rural categories were also shown to have some effect on ALL incidence in this dataset (Tables 7.22 and 7.24), and have been occasionally identified by the literature as important (but sometimes with contradictory results). Consequently, these were then also added separately to the base model to see if their inclusion had any effect on the associations between ALL and the various population mixing measures after age group and sex had been controlled for. The final models are shown in Figure 7.20 and these regression models were used as a template for all of the time periods analysed (1980-2004; 1980-1991 and 1992-2004; 1980-1986, 1987-1992, 1993-1998 and 1999-2004). Therefore, the control variables which had a significant effect on ALL

incidence in the 12/13 or 25 year study periods were included in the 6/7 year multivariate models even where no significant effect was found in the 6/7 univariate results. These variables were included since they have all been linked with ALL in the literature (see chapter 3) and shown by the 12/13 or 25 year analyses to be important in the New Zealand setting.

Table 7.24: Comparison of ALL univariate control models to the null model 1980-2004

Model	Deviance	Reduction in Deviance
Null	5,647	-
Age group	5,118	-529
Sex	5,245	-402
< 3 household members	5,247	-400
> 6 household members	5,250	-397
Urban/Rural (7 Categories)	5,252	-395
Population Density	5,255	-392
% Europeans	5,256	-391
Deprivation Score	5,256	-391
Urban/Rural (2 Categories)	5,257	-390

Figure 7.20: ALL regression models

Regression models:

Model 1:	Univariate	Population mixing measure (+ age- + sex-specific population at risk as an exposure variable)
Model 2:	Base model	Age group at diagnosis + sex + population mixing measure (+ age- + sex-specific population at risk as an exposure variable)
Model 3:	Alternative	Base model + % European
Model 4:	Alternative	Base model + deprivation score
Model 5:	Alternative	Base model + population density
Model 6:	Alternative	Base model + % households with less than 3 usual members
Model 7:	Alternative	Base model + % households with more than 6 usual members
Model 8:	Alternative	Base model + urban/rural (7 categories)
Model 9:	Alternative	Base model + urban/rural (2 categories)

7.4.2.3 Multivariate model results

Incorporating control variables into the models did not substantially alter the associations between ALL and the population mixing change variables for the 12/13/25 year analyses (Table 7.25). For example, the IRRs for the population change variable remained between 0.998 and 0.999 after controlling for each of the potential confounding variables, during all of the analysis periods. Furthermore, the confidence intervals for these IRRs varied little and all included one.

The inclusion of control variables did not affect the positive association noted between ALL and the change in the percentage of total migrants' variable between 1980 and 1991. The IRR of 1.003 (CI = 1.000-1.006) was observed in models 2-5 and suggests a significantly raised incidence of ALL in areas which increased the most in the percentage of migrants during this time. Adding either the household overcrowding or urban/rural variables (models 6-9) reduced the IRRs slightly to 1.002 and widened the confidence intervals. During the next 13 years (1992-2004) the IRRs for this variable were below one in every model, but were not statistically significant. During the same time period, the change in the percentage of child migrants' variable was also inversely associated with the count of ALL cases. This association was significant at the 90 percent confidence level in the univariate analysis (IRR = 0.996, CI = 0.992-1.000) and became significant at the 95 percent confidence level after the inclusion of age, sex and population density in the models (2 and 5) (IRR for model 5 = 0.995, CI = 0.990-0.999). Thus ALL was significantly lower in areas which had increased the most in the percentage of child migrants in the latter half of the study. The IRRs for this variable in the 12 and 25 year analysis were consistently above one although the confidence intervals all included one.

The associations noted between ALL and the change in the percentage of overseas visitors and migrants, and migrant diversity variables in the univariate models (Table 7.25) were not affected by the addition of the control variables for any time period. The IRRs for these variables remained very close to unity in all of the multivariate models (Table 7.25).

ALL continued to be significantly raised in category 2 areas (no change in population mixing category) compared to category 1 areas (decrease in population mixing category) between 1980 and 1991 after including age group, sex, deprivation score and population density in the models (models 2, 4, and 5). This association was almost significant in the latter half of the study period as well (IRR for model 5 = 1.258, CI = 0.938-1.686). However, the confidence intervals in all of the models included one. Also consistent with the univariate model was the raised ALL IRR for category 4 areas (increase in migration category) compared to category 1 areas for the period 1980-2004 (IRR for model 5 = 1.113, CI = 0.915-1.352). The inclusion of control variables decreased the IRR slightly and widened the confidence intervals which all included one.

Table 7.25: Results of the ALL multivariate regression models by year group and population mixing change variable

Population mixing change variable	1980-1991 Poisson	1992-2004 Negative binomial	1980-2004 Negative binomial
Population change	0.998 (0.995-1.002) SIG: NONE	0.999 (0.998-1.001) SIG: NONE	0.999 (0.999-1.000) SIG: NONE
Change in % total migrants	1.003* (1.000-1.006) SIG: M2-5 (+)	0.995 (0.987-1.003) SIG: NONE	1.001 (0.999-1.002) SIG: NONE
Change in % child migrants	1.002 (0.999-1.005) SIG: NONE	0.995* (0.990-0.999) SIG: M2,5 (-)	1.000 (0.998-1.002) SIG: NONE
Change in % overseas migrants	1.001 (0.999-1.002) SIG: NONE	1.000 (0.998-1.002) SIG: NONE	1.000 (0.999-1.000) SIG: NONE
Change in % overseas visitors	0.999 (0.999-1.000) SIG: NONE	0.999 (0.999-1.000) SIG: NONE	0.999 (0.999-1.000) SIG: NONE
Change in migrant diversity	0.997 (0.988-1.006) SIG: NONE	0.999 (0.990-1.007) SIG: NONE	1.001 (0.998-1.003) SIG: NONE
Population mixing change category:			
1. Decrease in either/both categories (base)	1.000	1.000	1.000
2. No change in either category	1.299* (1.007-1.674) SIG: M2,4,5 (+)	1.258 (0.938-1.686) SIG: NONE	1.053 (0.877-1.264) SIG: NONE
3. Increase in diversity category	1.276 (0.858-1.897) SIG: NONE	0.795 (0.446-1.417) SIG: NONE	1.029 (0.792-1.335) SIG: NONE
4. Increase in migration category	0.989 (0.722-1.354) SIG: NONE	1.051 (0.717-1.539) SIG: NONE	1.113 (0.915-1.352) SIG: NONE

IRRs are reported for population mixing (PM) measures in model 5 (age, sex & population density). * denotes a p-value of <0.05. 95% confidence intervals are reported in the parentheses. SIG denotes significant models, and the direction of the PM association in these models (+ or -) is noted in the parentheses.

Dividing the study period further into 6/7 year time periods (Table 7.26) revealed some trends which were consistent with the 12/13/25 year multivariate analyses (Table 7.25). For example, the majority of IRRs for the population change variable were less than one in every 6/7 year period. As with the 12/13 year analyses, the confidence intervals all included one. However, there was one exception. During the period 1993-1998, population change was positively associated with childhood ALL after controlling for the percentage of households with less than three usual members (IRR = 1.008, CI = 1.001-1.015). Thus, an increase in the total population was associated with a significantly raised risk of ALL in one model for this period.

Also consistent with the 12/13 year analyses was a significant association observed between ALL and the change in the percentage of total migrants in the earliest time period (1980-1986). During this time, raised ALL incidence was noted in areas which had increased the most in this population mixing measure (IRR for model 5 = 1.008, CI = 1.003-1.012). This relationship was significant at the 95 percent confidence level in all of the models tested, and the effect size was slightly larger than in the 12 year analysis (IRR for model 5 in 1980-1991 = 1.003, CI = 1.000-1.006). However, in the following six years (1987-1992), the IRRs for this variable fluctuated around one and the confidence intervals included one. Between 1993 and 1998 this variable had a significant IRR of 1.023 (CI = 1.001-1.046) in model 6 only. In the final six years of the study the IRRs for this variable were below one in every model but were not significant.

The change in the percentage of child migrants' variable was also positively associated with ALL in every model between 1980 and 1986 (IRR for model 5 = 1.005, CI = 1.002-1.009). Again the IRRs were slightly larger than noted in the 12 year period between 1980 and 1991, and none of the confidence intervals included one in the seven year analysis. The IRRs for this variable were consistently below one in all of the following six year periods, but all had confidence intervals which included one.

In the 6/7 year multivariate models, IRRs for the change in migrant diversity variable varied around unity, and as with the 12/13 year models, were not statistically significant in any of the models. However, when changes in migrant diversity were examined in conjunction with changes in migration in the categorical population mixing change variable, some consistent trends were observed. For instance, the IRRs for category 2 areas (no change in either diversity or migration category) were greater than one after controlling for each of the potential confounding variables in every time period. However, while the IRRs for this category were significant in the 1980-1991 multivariate analyses, the confidence intervals for the IRRs in the 6/7 year analyses all included one. Raised IRRs were generally noted in category 3 areas (increase in diversity category), especially in the middle two time periods (1987-1992 and 1993-1998). When modelled in conjunction with age group, sex and the percentage of Europeans, the IRR for this category (4.207, CI = 1.422-12.447) implied that over four times as many ALL cases were diagnosed in these areas compared to the base category areas during the period 1993-1998. Category 4 areas (increase in migration category) had IRRs greater than one for every 6/7 time period examined, with especially high IRRs noted when population density was added to the base model between 1987 and 1992 (IRR = 3.283, CI = 0.986-10.93).

The 6/7 year analysis also identified three additional population mixing variables as being significantly related to childhood ALL. During the periods 1987-1992 and 1993-1998, the incidence of ALL was found to be higher in areas which had increased the most in overseas migrants, even after adjusting for the majority of the control variables (IRR for model 5 1987-1992 = 1.005, CI = 1.002-1.008). However, the IRRs for the remaining time periods were mostly below one with confidence intervals that included one. A slightly weaker positive association was also noted in 1987-1992 with the change in the percentage of overseas visitors' variable, but only after including age group, sex and population density in the model (IRR = 1.0004, CI = 1.000-1.001). The only significant link between ALL and population mixing noted between 1999 and 2004 was for the change in one year mobility variable. After controlling for each of the potential confounding variables, childhood ALL was significantly higher in areas where the one year mobility percentage increased the most (IRR for model 5 = 1.007, CI = 1.001-1.014).

Thus, the majority of the significant associations between ALL and population mixing were in a positive direction for both the 12/13/25 and 6/7 year multivariate analyses. This finding suggests that ALL incidence was higher in areas where population mixing had increased the most. Furthermore, these significant positive associations were noted between ALL and a range of different population mixing measures. There was only one significant exception: in the 13 years between 1992 and 2004, when increases in the percentage of child migrants were associated with a significant decrease in ALL in models 2 and 5.

Table 7.26: Results of the ALL multivariate regression models by 6/7 year group and population mixing change variable

Population mixing change variable	1980-1986 (Poisson)	1987-1992 (ZIP)	1993-1998 (ZIP)	1999-2004 (Poisson)
Population change	0.989 (0.978-1.001) SIG: NONE	0.994 (0.978-1.011) SIG: NONE	0.998 (0.990-1.006) SIG: M6 (+)	0.997 (0.991-1.003) SIG: NONE
Change in % total migrants	1.008*(1.003-1.012) SIG: ALL (+)	0.993 (0.979-1.007) SIG: NONE	1.001 (0.984-1.018) SIG: M6 (+)	0.991 (0.977-1.004) SIG: NONE
Change in % child migrants	1.005*(1.002-1.009) SIG: ALL (+)	0.991 (0.983-1.000) SIG: NONE	0.994 (0.985-1.004) SIG: NONE	0.994 (0.987-1.001) SIG: NONE
Change in % overseas migrants	0.999 (0.996-1.001) SIG: NONE	1.005*(1.002-1.008) SIG: M3-M9 (+)	1.001 (0.999-1.003) SIG: M2,4,6 (+)	0.999 (0.995-1.002) SIG: NONE
Change in % overseas visitors	1.000 (0.999-1.000) SIG: NONE	1.0004*(1.000-1.001) SIG: M5 (+)	0.999 (0.994-1.005) SIG: NONE	1.000 (0.999- 1.001) SIG: NONE
Change in 1 year mobility	NO DATA	0.995 (0.987-1.003) SIG: NONE	1.013 (0.993-1.033) SIG: NONE	1.007*(1.001-1.014) SIG: ALL (+)
Change in migrant diversity	0.999 (0.987-1.011) SIG: NONE	0.973 (0.945-1.003) SIG: NONE	0.994 (0.968-1.021) SIG: NONE	1.000 (0.985-1.015) SIG: NONE
Change in population mixing category :				
1. Decrease in either/both categories (base)	1.000	1.000	1.000	1.000
2. No change in either category	1.278 (0.917-1.780) SIG: NONE	1.282 (0.705-2.331) SIG: NONE	1.251 (0.760-2.061) SIG: NONE	1.345 (0.869-2.083) SIG: NONE
3. Increase in diversity category	0.832 (0.411-1.684) SIG: NONE	2.533 (0.802-7.999) SIG: NONE	1.719 (0.708-4.172) SIG: M3 (+)	0.505 (0.192-1.329) SIG: NONE
4. Increase in migration category	1.008 (0.629-1.615) SIG: NONE	3.283*(0.986-10.93) SIG: M5 (+)	1.281 (0.716-2.289) SIG: NONE	1.030 (0.595-1.781) SIG: NONE

IRRs are reported for population mixing measures in model 5 (age, sex & population density). * denotes a p-value of <0.05. 95% confidence intervals are reported in the parentheses. SIG denotes significant models, and the direction of the PM association in these models (+ or -) is noted in the parentheses.

7.5 Conclusions

This chapter has uncovered a number of important themes in the geographical epidemiology of childhood ALL in New Zealand between 1980 and 2004. First, while childhood ALL remains a relatively rare disease in New Zealand, its incidence appears to be increasing. Second, ALL has been shown to vary considerably by a number of individual characteristics, including, age group at diagnosis, sex and ethnicity. Third, childhood ALL displayed distinct spatial variation across the country, at a number of geographical scales. In particular, incidence was higher in the main urban centres of the country and several significant space-time clusters of this disease were identified in these areas. In addition, there was a reduced risk of childhood ALL in the most deprived areas of New Zealand.

Finally, the chapter has described the results of the first systematic analysis of the association between population mixing and childhood ALL for small areas across the whole of New Zealand. Multivariate regression analyses have shown ALL to be significantly related to a number of different measures of population mixing change in this setting. Generally, areas which increased the most in population mixing were found to have the highest number of ALL cases. This finding was especially apparent in the 6/7 year analyses, where a total of seven different population mixing change measures were found to be significantly and positively associated with ALL. Only one contradictory result was noted and occurred in the 13 year analysis: a significantly reduced risk of ALL was found in areas which had increased the most in child migrants between 1992 and 2004. However, this association was only significant in two models. Interestingly, the associations between ALL and each population mixing measure did not remain constant over time. The following chapter seeks to determine whether population mixing is also significantly associated with childhood type 1 diabetes in New Zealand.

Chapter 8: Population mixing and the geographical epidemiology of childhood type 1 diabetes in Canterbury, New Zealand

8.1 Introduction

The preceding chapter detailed the geographical epidemiology of childhood ALL in New Zealand and explored the possible associations between this disease and population mixing. The results showed how ALL varied over time and space, between areas of differing socioeconomic status, differing levels of household overcrowding, and in urban and rural New Zealand. The chapter also documented a statistically significant association between ALL and a number of measures of area-level population mixing. This chapter examines whether population mixing has a similar influence on the occurrence of childhood type 1 diabetes in the Canterbury region of New Zealand. First, the chapter considers the general epidemiology of this disease in children aged 0-14 years living within five district council areas in the region of interest. It explores trends by age, sex, ethnicity, year of diagnosis, and CAU of residence at diagnosis. Second, a number of methods are utilised to test the relationship between type 1 diabetes and population mixing at the CAU-level for different temporal periods.

8.2 Descriptive patterns

8.2.1 Individual-level

As with ALL, childhood onset type 1 diabetes has been shown to vary by a number of individual characteristics, including age, sex, and ethnicity. This section considers how children who present with type 1 diabetes in the Canterbury region differ in terms of these individual attributes, and begins with an overview of the disease in this region.

During the study period (1980-2004) a total of 337 new cases of type 1 diabetes were diagnosed in children aged less than 15 years in Canterbury. Of this total, 164 (48.7 percent) were male and 173 (51.3 percent) were female. The average age at diagnosis was similar for males (8.72 years) and females (8.86 years). The median age at diagnosis was 9.56 years for males and 9.73 years for females. The modal age at diagnosis was higher for males (13 years) than females (11 years) (Table 8.1).

Table 8.1: Type 1 diabetes descriptive statistics by age and sex

Descriptive Statistics	Male	Female	Total
Count	164	173	337
% of total	48.7	51.3	100
Mean age at diagnosis	8.72	8.86	8.79
Modal age at diagnosis	13	11	11
Median age at diagnosis	9.56	9.73	9.62
Standard deviation of age at diagnosis	4.22	4.15	4.18

Closer examination of cases by age at diagnosis and sex revealed that, as age increased, the number of type 1 diabetes cases diagnosed also tended to increase, with the majority of cases diagnosed in older children. The smallest number of cases diagnosed occurred in children aged less than one year (two cases; one male, one female). There was little variation in the age-specific count between males and females, although the number of female cases peaked slightly earlier (age 11) than male cases (age 13) (Figure 8.1). After adjusting for the population at risk, a clear trend by age group was evident (Figure 8.2). In both sexes the incidence per 100,000 population was lowest in the youngest age group (0-4 years) and increased with age. Female incidence rates were higher than male incidence in both the youngest and oldest age groups. The difference between the sexes was greatest in the oldest age group (616 male cases per 100,000 population compared to 708 females cases) and was smallest in the middle age group (450 male cases compared to 431 females cases).

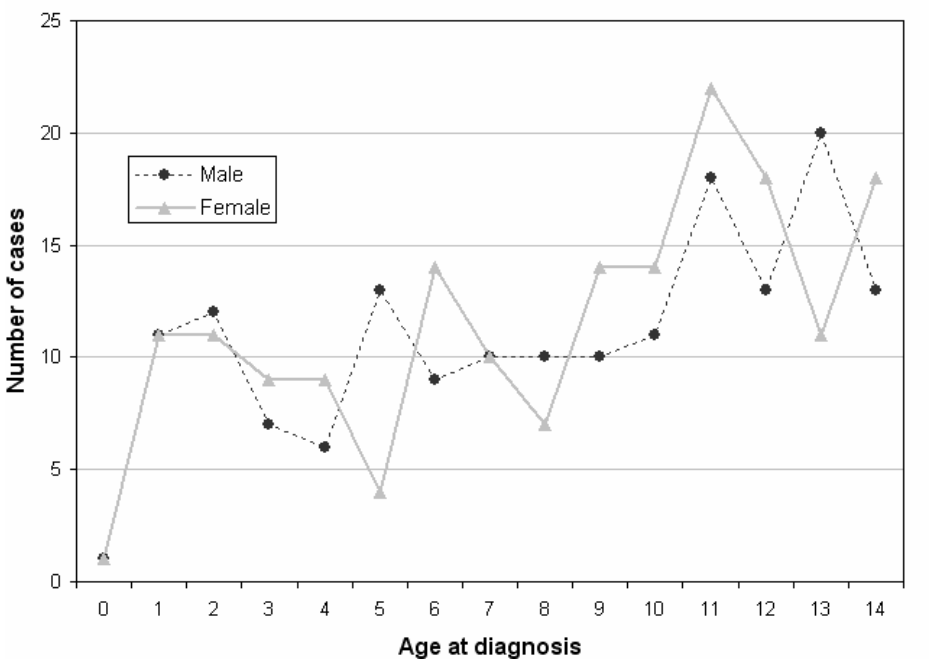

Figure 8.1: Number of type 1 diabetes cases in Canterbury by age at diagnosis and sex, 1980-2004

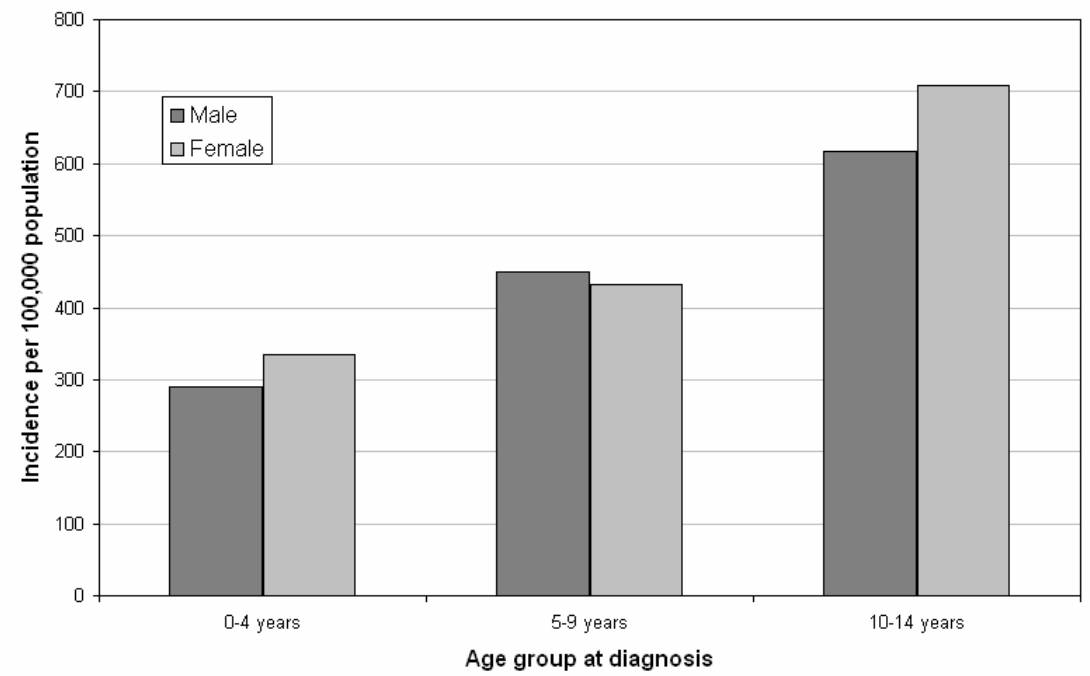

Figure 8.2: Type 1 diabetes incidence in Canterbury per 100,000 population
by age group and sex, 1980-2004

Appreciable differences in incidence were also noted by ethnic group. 84.57 percent of all children diagnosed with type 1 diabetes in Canterbury between 1980 and 2004 were of New Zealand European ethnicity. The next highest incidence occurred in children coded as 'other' who represented 9.20 percent of the total cases, followed by those whose ethnic group was not stated (2.97 percent) and those classed as 'other European' (1.48 percent). New Zealand Māori, Chinese and other Pacific each represented less than 1 percent of the total cases registered over the 25 year study period (Table 8.2).

Incidence rates were calculated in order to account for differences in the size of the child populations by ethnic group within the Canterbury region (Figure 8.3). New Zealand European and Other European children were grouped together and those falling into the 'other' and 'not stated' categories were not included. Even after controlling for the population at risk, the incidence rates for European children were considerably higher than those for Māori, Pacific and Chinese children. During the 25 year study period there were approximately 414 new cases of type 1 diabetes diagnosed per 100,000 population in European children, compared to 51 per 100,000 in Chinese children, 33 per 100,000 in Pacific children, and 30 per 100,000 in Māori children. Due to the small number of total cases diagnosed in the Māori (3), Pacific (2) and Chinese children (1), analysis of incidence trends by ethnicity and age group was not possible.

Table 8.2: Type 1 diabetes counts for Canterbury 1980-2004 by ethnic group, age group at diagnosis and sex

Ethnic Group	Sex	0-4 years	5-9 years	10-14 years	All ages	% of total cases
	Male	33	46	68	147	43.62
New Zealand	Female	34	40	64	138	40.95
European	**Total**	**67**	**86**	**132**	**285**	**84.57**
	Male	3	3	6	12	3.56
Other	Female	3	1	15	19	5.64
	Total	**6**	**4**	**21**	**31**	**9.20**
	Male	1	2	0	3	0.89
Not Stated	Female	0	5	2	7	2.08
	Total	**1**	**7**	**2**	**10**	**2.97**
	Male	0	0	1	1	0.30
Other	Female	1	2	1	4	1.19
European	**Total**	**1**	**2**	**2**	**5**	**1.48**
	Male	0	0	0	0	0.00
New Zealand	Female	2	0	1	3	0.89
Māori	**Total**	**2**	**0**	**1**	**3**	**0.89**
	Male	0	1	0	1	0.30
Chinese	Female	0	1	0	1	0.30
	Total	**0**	**2**	**0**	**2**	**0.59**
	Male	0	0	0	0	0.00
Pacific	Female	1	0	0	1	0.30
	Total	**1**	**0**	**0**	**1**	**0.30**
Total		**78**	**101**	**158**	**337**	**100**

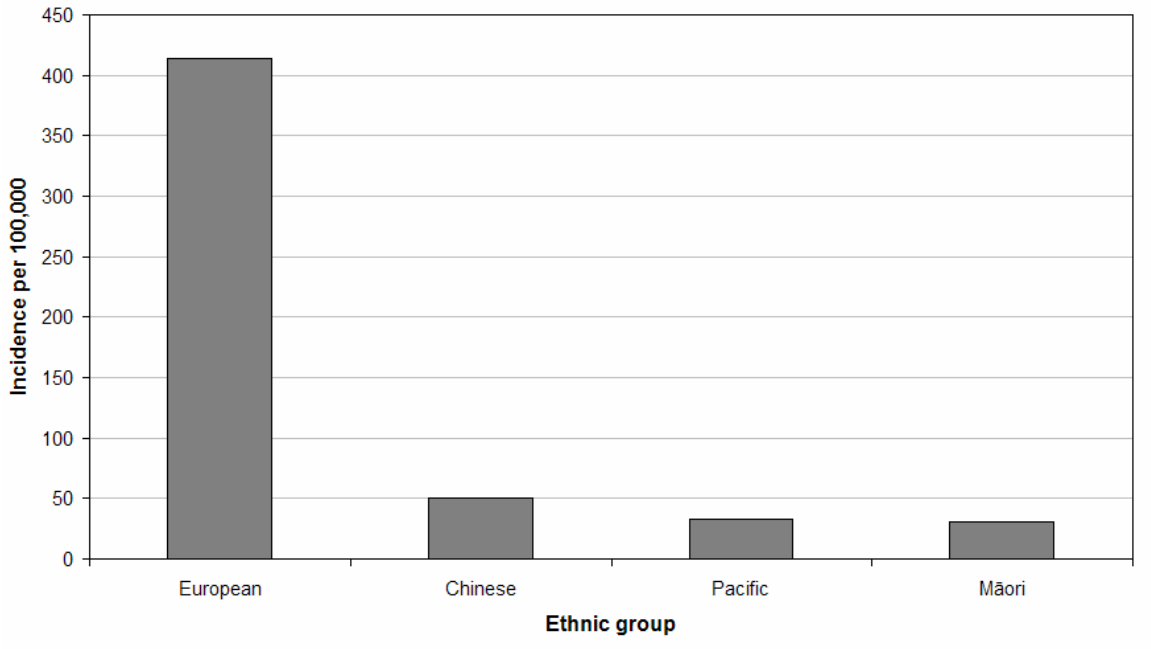

Figure 8.3: Type 1 diabetes incidence in Canterbury by ethnicity, 1980-2004

There was an overall increase in incidence of type 1 diabetes over the 25 year period (Figure 8.4). In 1980 there were 5.3 cases of type 1 diabetes diagnosed per 100,000 population but by

2004, this rate had risen to 32.5 cases per 100,000. Prior to 1989, the incidence rate only rose above 15 cases per 100,000 once (in 1984) and was on average 10 cases per 100,000. After 1989 the incidence rate only dropped below 15 cases per 100,000 once (in 1993) and for the majority of the time remained above 20 cases per 100,000. Since 1998 the incidence was frequently greater than 25 cases per 100,000. The greatest increase in incidence occurred between 2003 and 2004 when the incidence rate more than doubled from 16.2 to 32.5 cases per 100,000.Using linear regression, the average increase in incidence was 0.86 cases per 100,000 per year.

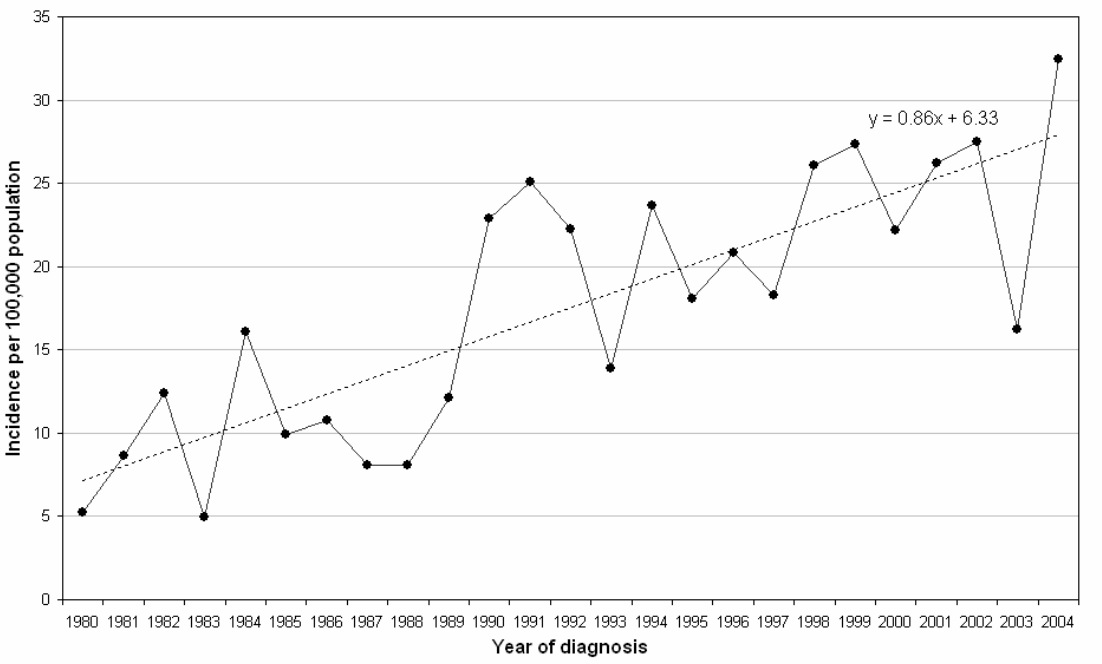

Figure 8.4: Incidence of type 1 diabetes in Canterbury, by year of diagnosis

Examining temporal trends by age group at diagnosis was problematic due to the small number of cases involved. As a result, incidence rates were calculated using three year moving averages to remove short term variation (Figure 8.5). The results show a general increase in incidence in all age groups over the 25 year study period. Incidence rates were highest in the 10-14 year age group for all years except 1991, where rates in the 5-9 year age group were slightly higher. The greatest increase in incidence for 10-14 year olds occurred between 1987 and 1990 when incidence rose from 14.12 to 24.95 cases per 100,000 population. After this time, incidence levelled out for three years and then increased steadily to its peak of 36.27 cases per 100,000 population in 2001. Even after smoothing the data, yearly peaks and troughs were noted after 1994. Between 1980 and 1988, incidence in the 0-4 and 5-9 year age groups was much lower than for the 10-14 years group, and was on average 6 cases per 100,000 population per year. However, after 1988 the incidence for both age groups increased considerably; to 22.13 and

24.18 cases per 100,000 population in 1991. Incidence in 0-4 year olds then fell sharply to 8.05 cases per 100,000 in 1994, with a more gradual decrease noted for 5-9 year olds. Incidence has generally been increasing again for the 0-4 year age group since 1994, and since 1996 for the 5-9 year age group.

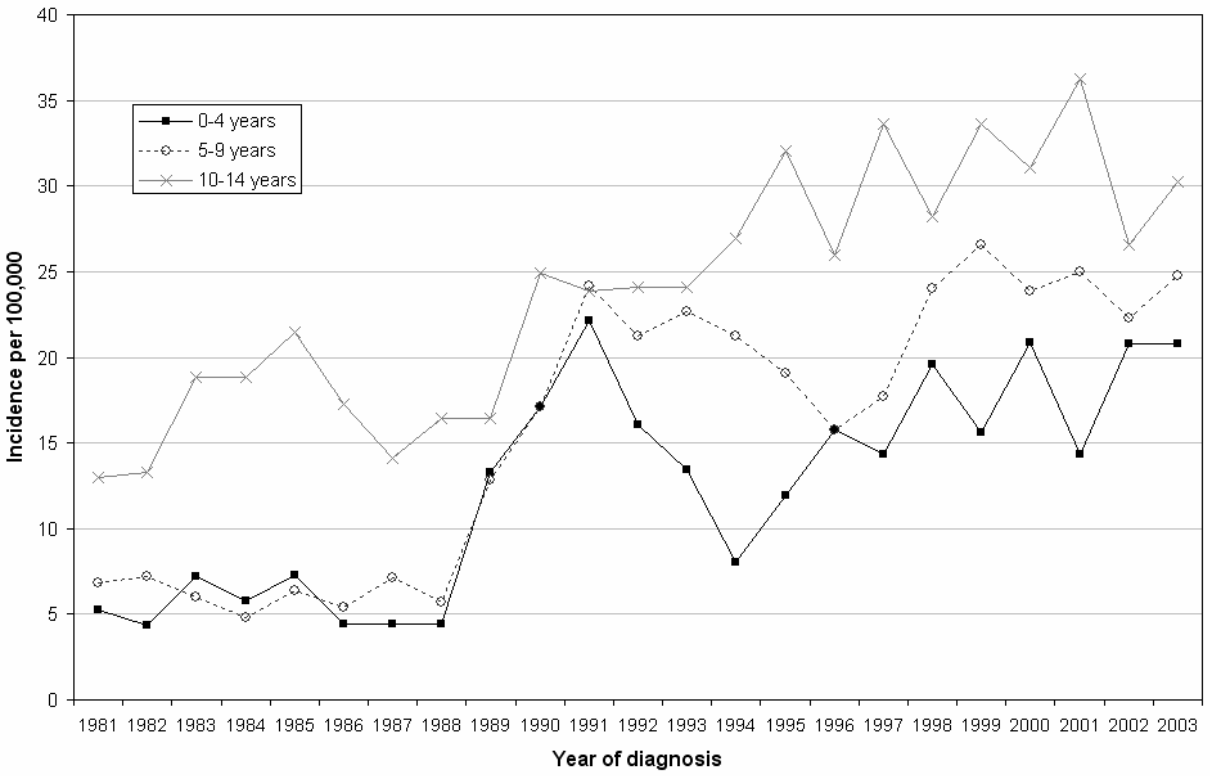

Figure 8.5: Incidence of type 1 diabetes in Canterbury, by year of diagnosis and age group
(3 year moving averages)

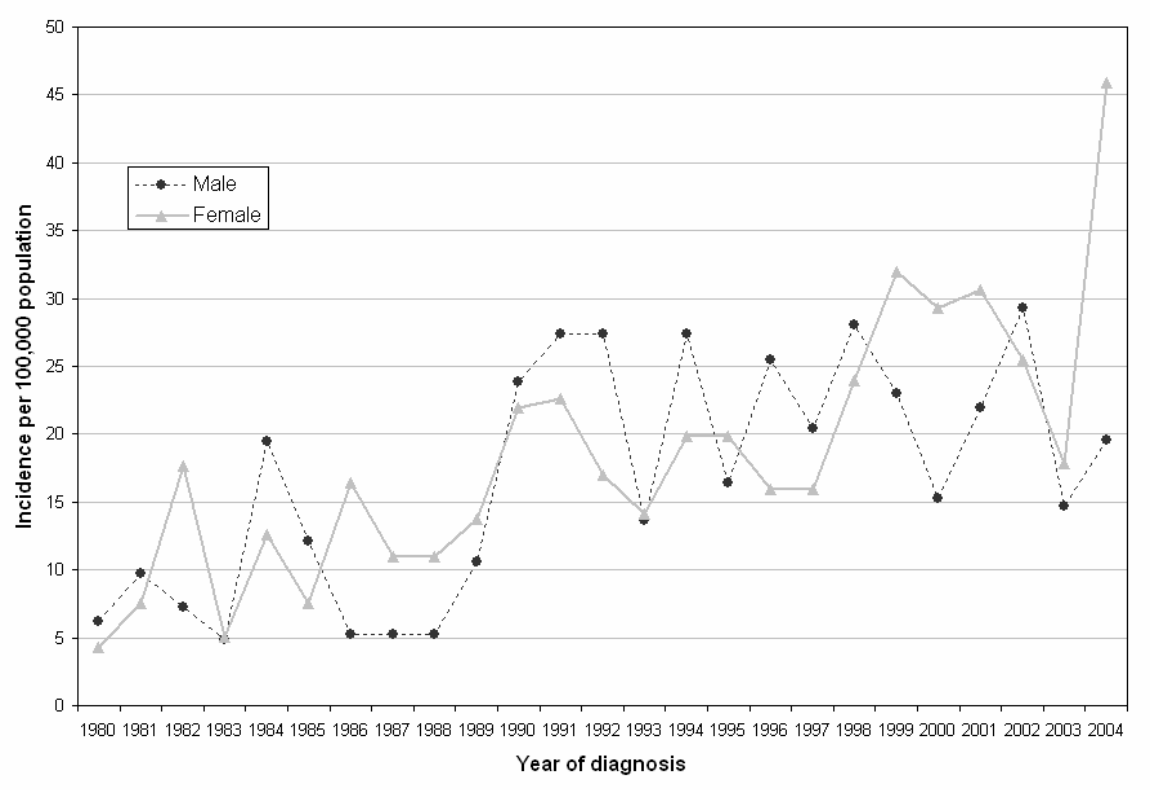

Figure 8.6: Incidence of type 1 diabetes in Canterbury, by year of diagnosis and sex

Trends by sex were slightly less variable than by age group. Male rates were higher than female rates in a total of 12 years and female rates were higher than male rates in the remaining 13 years (Figure 8.6). The rates fluctuated from being higher in males than females on a one-to-four year timescale. Both the lowest (4.3 cases per 100,000 in 1980) and highest (45.9 cases per 100,000 in 2004) rates were noted in females. The lowest male incidence of 4.9 cases per 100,000 occurred in 1983 and the highest male incidence was 29.3 cases per 100,000 in 2002. The largest difference between male and female incidence was recorded in 2004 when around 19.5 male cases were diagnosed per 100,000 population compared with 45.9 female cases per 100,000.

8.2.2 Area-level

In addition to the individual-level characteristics, childhood type 1 diabetes has also been associated with a range of area-level characteristics, including measures of social deprivation (Crow et al., 1991, Parslow et al., 2001). Using the New Zealand Deprivation Score for 2001 (NZDep 2001), observed and expected cases were summed by deprivation deciles and standardised incidence ratios (SIRs) were calculated.

There was an approximately linear trend in the rates of type 1 diabetes between the least deprived tenth of neighbourhoods (decile 1) and the most deprived neighbourhoods (decile 10) with the highest SIR (120.50) in decile 1 and the lowest SIR (70.30) in decile 10 (Table 8.3). The SIRs of the four least deprived deciles (1-4) were all greater than 100 indicating that type 1 diabetes was especially high in these areas compared with the regional average. The more deprived deciles 5,6,7,8 and 10 all had SIRs below 100 signifying that fewer type 1 diabetes cases than expected were noted in these areas during the study period (Figure 8.7). One exception to this trend was the SIR of 105.72 in deprivation decile 9. However, the 95 percent confidence intervals (CIs) for this SIR demonstrate that its value could be anywhere between 55.50 and 155.94. It should be noted that none of the SIRs were statistically significant according to their chi-square values and they each had wide CIs that included 100 (Table 8.3).

Table 8.3: Type 1 diabetes SIRs, chi-square values and CIs by deprivation decile for Canterbury, 1980-2004

Deprivation Decile	Observed	Expected*	SIR	Lower CI	Upper CI	Chi-square	Significant
1 (Low)	48	39.84	120.50	83.08	157.91	1.67	NO
2	47	40.02	117.45	81.06	153.84	1.22	NO
3	31	29.36	105.60	67.40	143.80	0.09	NO
4	55	51.94	105.89	77.09	134.69	0.18	NO
5	44	47.31	93.00	66.50	119.51	0.23	NO
6	27	27.55	98.02	61.41	134.62	0.01	NO
7	24	30.31	79.18	50.99	107.37	1.31	NO
8	34	40.86	83.22	57.70	108.74	1.15	NO
9	18	17.03	105.72	55.50	155.94	0.06	NO
10 (High)	9	12.80	70.30	31.79	108.81	1.13	NO

*Expected cases = sum of (age-specific population in each decile x age-specific incidence rate)

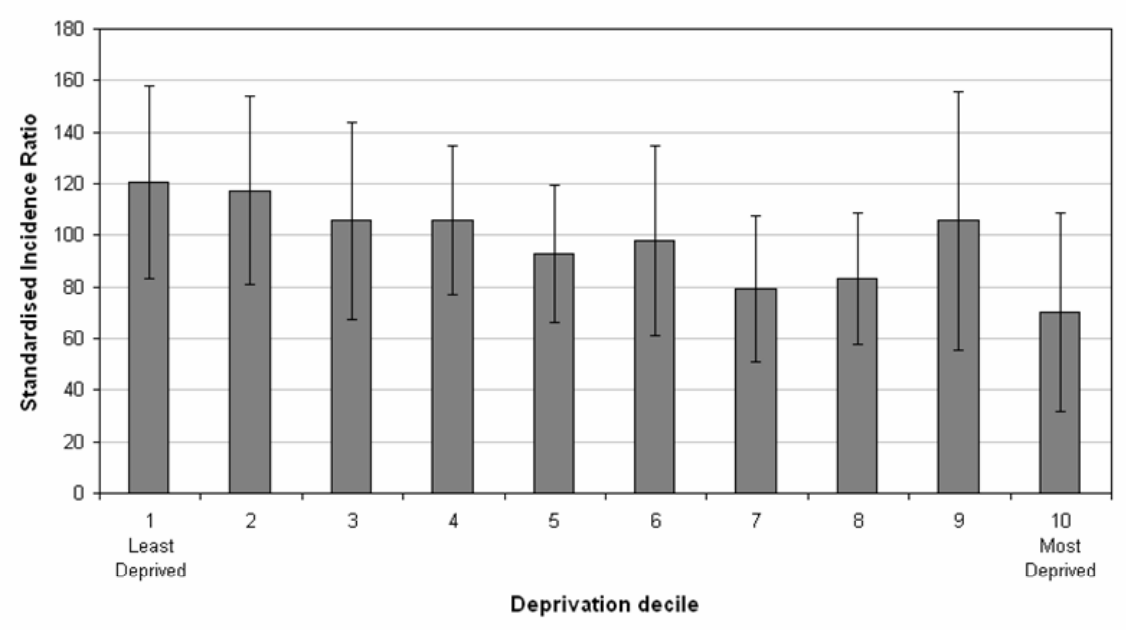

Figure 8.7: Type 1 diabetes SIRs by deprivation decile for Canterbury, 1980-2004
(Bars represent 95% CIs)

When areas were aggregated further to account for the small number of cases observed in each decile of neighbourhoods, a similar pattern in incidence by deprivation quintile was observed. As deprivation increased, rates of type 1 diabetes tended to decrease (Figure 8.8). The least deprived areas of Canterbury (quintile 1) recorded the highest SIR of 118.97 (CI = 92.88-145.06). The two most deprived quintiles (4 and 5) both had SIRs of less than 100 with upper CIs of 100.43 and 123.00 respectively. These figures imply that type 1 diabetes was higher in the more affluent areas and lower in the more deprived areas when compared to the regional average. The CIs for these SIRs were much narrower than those produced by the decile-level analysis but they all still included 100 and none of the chi-square values were significant.

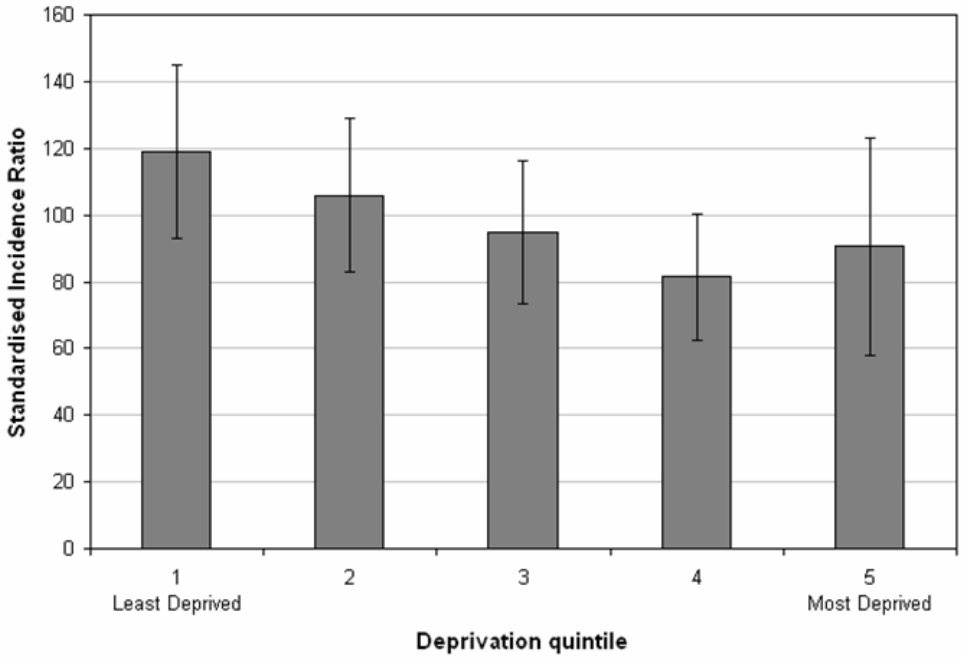

Figure 8.8: Type 1 diabetes SIRs by deprivation quintile for Canterbury, 1980-2004
(Bars represent 95% CIs)

In order to determine whether there were urban/rural variations in type 1 diabetes, SIRs were calculated for seven urban/rural categories, ranging from main urban area to highly rural/remote area (Table 8.4). There was no trend in incidence across the urban/rural spectrum. However, SIRs below 100 were noted in the three most rural categories. No cases of type 1 diabetes were observed during the study period in areas classed as independent urban communities (0.49 cases were expected) and rural areas with a low urban influence (0.72 cases were expected). An excess of cases was noted in areas categorised as satellite urban communities (SIR = 162.77) and rural areas with a high urban influence (SIR = 138.09) although the error margins for these SIRs were large. The SIR of 162.77 for satellite urban communities was considered statistically significant according to its chi-square value, and had a lower CI of 87.19 and an upper CI of 238.35. Main urban areas experienced around six fewer cases than were expected according to the age-specific populations at risk and the age-specific incidence rates for the whole of Canterbury (Table 8.4).

To account for the small numbers of cases in some of the categories, areas within the region were reclassified as either being predominantly urban or predominantly rural. The SIR for predominantly urban areas (101.50) was slightly greater than the regional average of 100 whereas the SIR for predominantly rural areas (88.69) was below average, suggesting that higher incidence of type 1 diabetes occurred in urban areas compared to rural areas over the study

period. Once again however, the chi-square values were not statistically significant and the true values of both SIRs could have been below, the same as, or above 100 (Table 8.5).

Table 8.4: Type 1 diabetes SIRs, chi-square values and CIs by urban/rural classification for Canterbury, 1980-2004

Urban/rural Category	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
Main Urban Area	273	279.23	97.77	86.30	109.24	0.14	NO
Satellite Urban Community	29	17.82	162.77	87.19	238.35	7.02	YES
Independent Urban Community	0	0.49	0.00	0.00	0.00	0.49	NO
Rural Area with High Urban Influence	15	10.86	138.09	55.97	220.21	1.58	NO
Rural Area with Moderate Urban Influence	12	18.33	65.46	35.49	95.43	2.19	NO
Rural Area with Low Urban Influence	0	0.72	0.00	0.00	0.00	0.72	NO
Highly Rural/ Remote Area	8	9.55	83.79	30.64	136.95	0.25	NO

Table 8.5: Type 1 diabetes SIRs, chi-square values and CIs by predominantly urban/rural area for Canterbury, 1980-2004

Urban/rural Category	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
Predominantly urban areas	302	297.54	101.50	89.97	113.03	0.07	NO
Predominantly rural areas	35	39.46	88.69	61.02	116.36	0.50	NO

8.3 Geographical distribution

Previous research has shown that even after controlling for urban/rural status and area-level deprivation, significant geographical variations in type 1 diabetes incidence remain (Patterson and Waugh, 1992). This finding indicates the existence of other important risk factors which can be explored by mapping incidence rates and examining the spatial patterns.

8.3.1 Geographical distribution of type 1 diabetes SIRs

In order to determine the geographical distribution of type 1 diabetes in children in the Canterbury region, SIRs were calculated at the census area unit (CAU)-level, for the periods 1980-2004, 1980-1991 and 1992-2004. Due to the small number of cases involved, confidence intervals and chi-square values for these ratios were also computed.

The map for the whole study period (1980-2004), shows two large areas with SIRs greater than 100 to the north and west of the region (Figure 8.9). In the north, the highly rural CAU of Amuri recorded a SIR of 104.1 suggesting an excess of type 1 diabetes cases in this area. However, this SIR was based on only 2 observed cases over the 25 year period where 1.9 cases were expected. As a result, its CIs were extremely wide (12.5-376.1) highlighting the instability of the ratio. Similarly, in the west of the region, the CAU of Malvern had a SIR of 102.0 based on 3 observed cases, with CIs of 21.1-298.3. Between these two areas, in the central belt in southern Hurunui, there was a band of CAUs whose SIRs were below the average for the region, as was the case in the south of the region in the CAUs of Kirwee and Selwyn-Rakaia. However their CIs included 100. A relatively large section in the middle of the region, in west of the Waimakariri District Council observed no new cases of type 1 diabetes in 0-14 year olds and therefore all had SIRs of zero. The highest SIR of 742.2 was noted in Riccarton South in the Christchurch main urban area and occurred as a result of 2 observed cases compared to 0.3 expected cases over the study period. This SIR had a relatively high lower confidence interval of 89.1 and a very high upper confidence interval of 2,679.3, but still included the regional average of 100. A further high SIR (502.6) was observed in Southbrook to the south of Rangiora in the Waimakariri District Council. The chi-square value for both these SIRs suggested that they were statistically significant.

Detailed examination of the main urban area of Christchurch showed that CAUs in the central city area all had SIRs of below 100. CAUs on the periphery of the central city all had SIRs of greater than 100. Large contiguous areas with SIRs which were above the regional average occurred to the north-east and south-east of the city, and also on its western fringe.

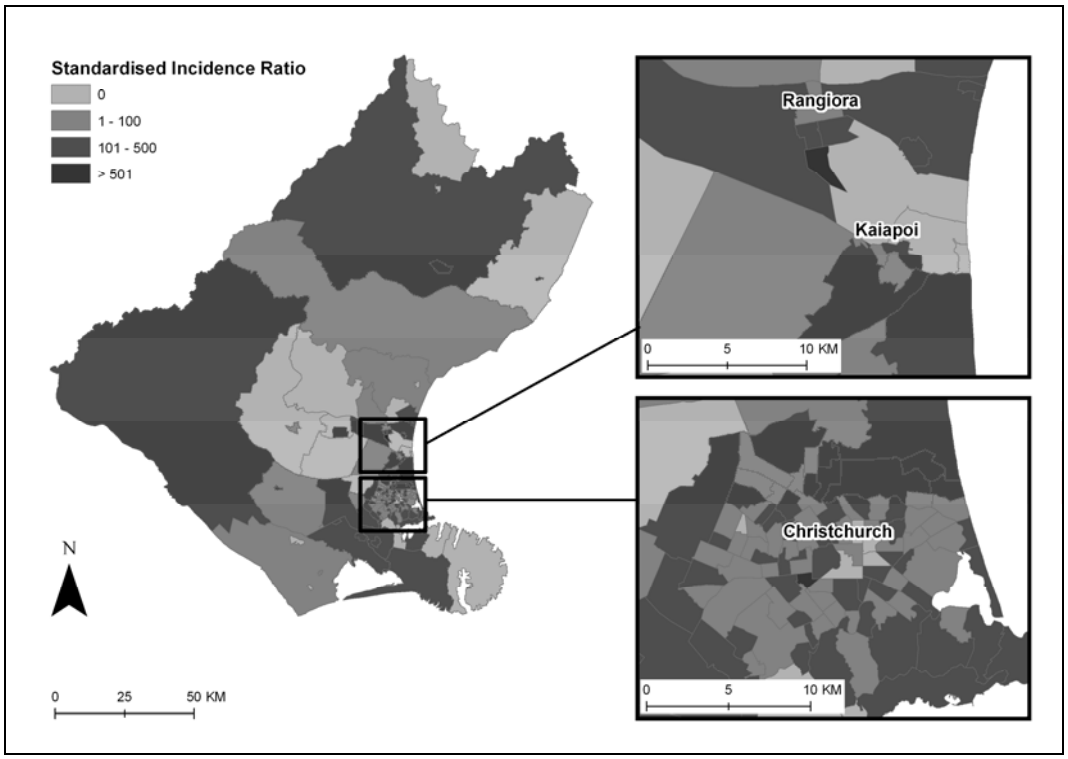

Figure 8.9: Type 1 diabetes SIRs in Canterbury 1980-2004

Dividing the study period into two approximately equal intervals allowed the examination of spatial trends in type 1 diabetes incidence over time. These analyses were especially important considering the substantial increase in incidence that has occurred over the 25 year study period (Figure 8.4). Between 1980 and 1991 the majority of rural Canterbury had SIRs of zero as no new cases were diagnosed in these areas during this time (Figure 8.10). Two notable exceptions included the highly rural/remote areas of Culverden and Cheviot in the Hurunui District Council. In Culverden a statistically significant SIR of 967.0 was recorded with CIs of 117.1-3,523.4 and a chi-square value of 15.7, implying that significantly more cases were observed than expected compared to the Canterbury average. More cases than expected were also observed in Cheviot which had a SIR of 679.3. However the CIs for this SIR included 100 and therefore this SIR value could have occurred by chance. Higher incidence than expected was also noted in the satellite urban community of Rangiora, the rural area with high urban influence of Waikuku, and the main urban area of North Kaiapoi during this period. A large contiguous area of high incidence occurred in the south of the study region in rural areas with a moderate to high urban influence from the neighbouring Christchurch area. Within the central Christchurch urban area, a patchwork of varying incidence ratios were noted with adjoining areas of high SIRs to the north-east, south-east and west of the city centre. Very high ratios were observed in the Middleton

(697.3) and Halswell West (1,215.0) CAUs in the south-west of Christchurch, although both figures were only based on one new diagnosis over the twelve year period.

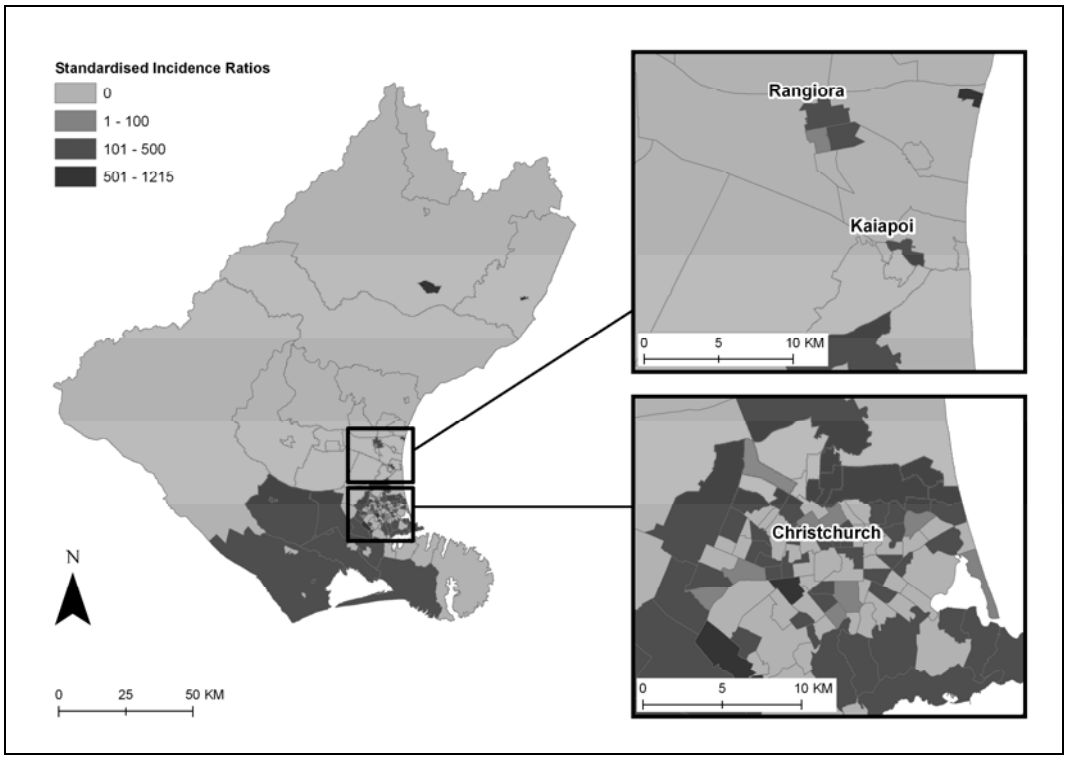

Figure 8.10: Type 1 diabetes SIRs in Canterbury 1980-1991

Compared to the earlier period (1980-1991), the latter half of the study (1992-2004) displayed pronounced differences in incidence (Figure 8.11). In particular, the rural areas to the north and west of the region which had SIRs of zero between 1980 and 1991, each recorded incidence ratios above 100 between 1992 and 2004. The CAUs to the south-west of Christchurch had SIRs of zero in the last thirteen years, compared to the higher than average ratios shown at the beginning of the study. CAUs in Rangiora, Waikuku and North Kaiapoi however all retained SIR values of over 100 for this period. Incidence ratios within Christchurch city centre displayed a similar distribution to that noted for the whole study period; zero or lower than average SIRs were found in the central city areas with an area of higher incidence CAUs around this. Incidence rates were also higher in CAUs to the south-east of the city centre. For the period 1992-2004, Riccarton South had the highest SIR of 1,056.76 (CI = 126.81-3,814.91) based on 2 observed cases where only 0.19 cases were expected. The SIR was statistically significant according to its chi-square value of 17.32. There were no cases observed in this CAU during the first 12 years of the study. Between 1992 and 2004, the CAU was surrounded by a number of CAUs with high SIRs to the north.

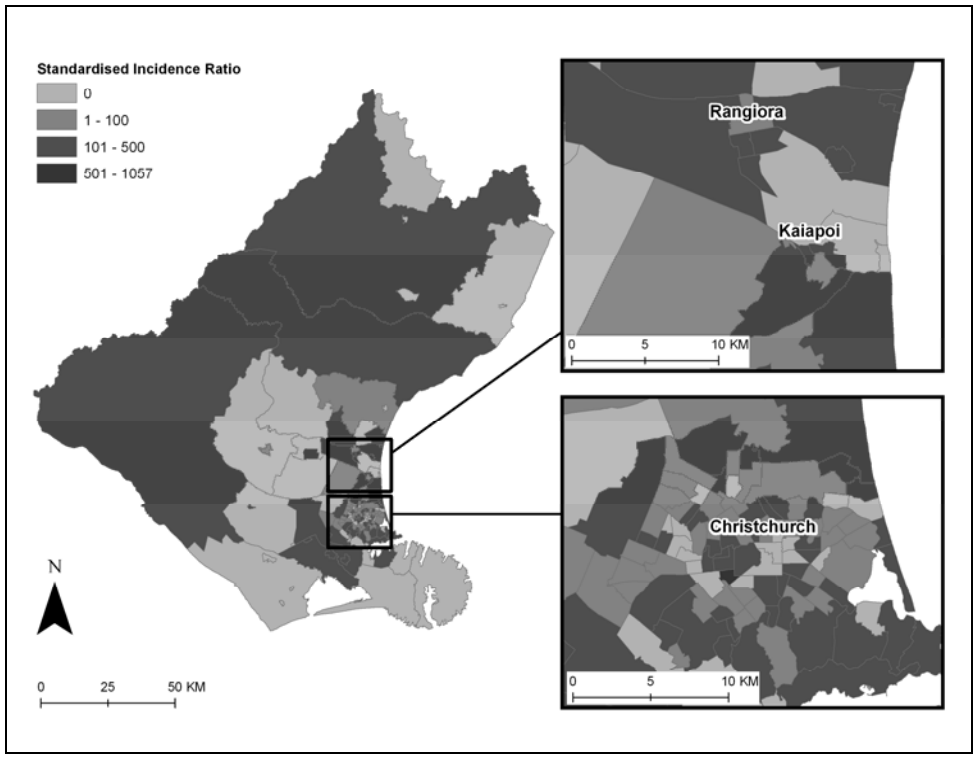

Figure 8.11: Type 1 diabetes SIRs in Canterbury 1992-2004

8.3.2 Geographical distribution of type 1 diabetes Poisson probabilities

The calculation of SIRs by small areas revealed distinct geographical patterns in type 1 diabetes incidence within the Canterbury region. However it is possible that the SIR values may be unreliable due to the small number of cases occurring in each CAU. Poisson probabilities were calculated to distinguish between areas of higher or lower than expected incidence. Low Poisson probabilities indicate an excess of observed cases, and areas whose Poisson probabilities were smaller than 0.05 represent places where there were significantly more cases of disease than would be expected by the Poisson distribution (at the 95 percent confidence level).

Over the entire study period the majority of CAUs in Canterbury had Poisson probabilities greater than 0.05 (Figure 8.12) which suggests that the observed numbers were within the range anticipated from a Poisson distribution. All of the significant excesses of cases occurred in the main urban area of Christchurch. Two were found in the north of the city centre in the suburbs of Redwood North (0.030) and Burwood (0.032); two to the south-east of the city in the adjoining areas of Moncks Bay (0.032) and Sumner (0.036); three to the south west of the city in Riccarton South (0.028), Hillmorton (0.021) and Oaklands (0.028) and one directly south of the city centre in Cashmere West (0.019). All of these areas were highly urbanised and six out of these eight

were classed as deprivation decile 1-4 areas (more affluent areas) and the remaining two were decile 7 and 8 areas (more deprived areas).

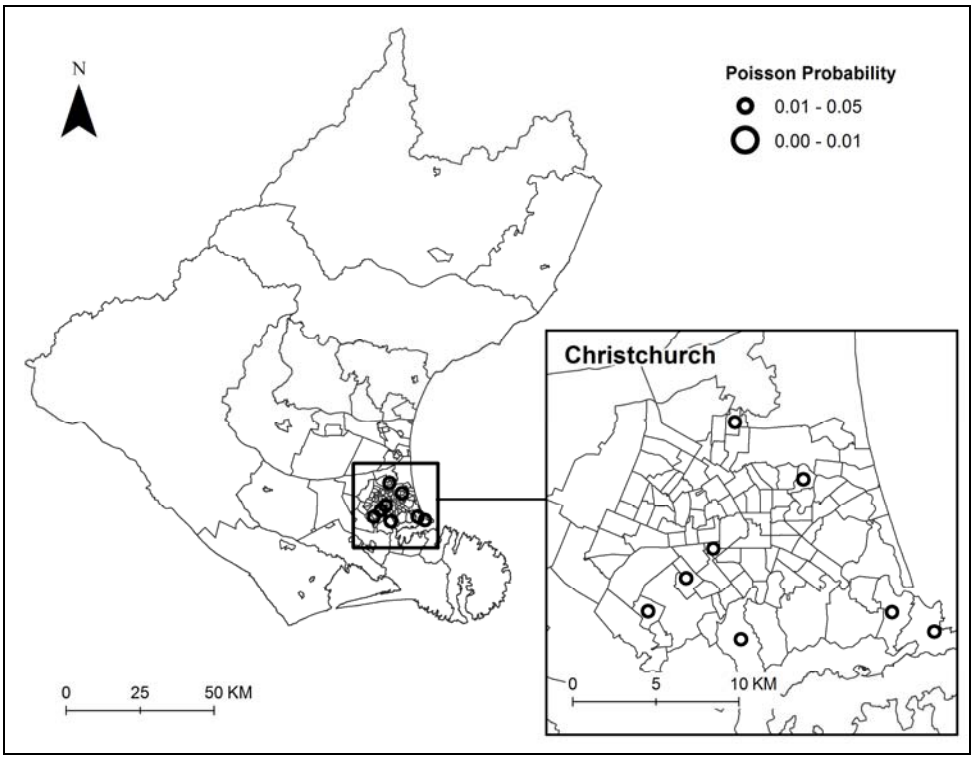

Figure 8.12: Type 1 diabetes Poisson probabilities 1980-2004

When the results where disaggregated into two time periods (1980-1991 and 1992-2004), the first period had only three CAUs with a significant excess of type 1 diabetes diagnoses (Figure 8.13). Two of these areas were situated within the main urban area of Christchurch; in the suburbs of Oaklands (0.045) to the south-west of the city centre and Moncks Bay (0.043) to the south-east. In addition, the highly rural/remote area of Culverden in the north of the region had more cases of type 1 diabetes (0.017) than were expected. During the later time period (1992-2004) only four CAUs, all within Christchurch, recorded Poisson probabilities of less than 0.05 (Figure 8.14). These CAUs included the suburbs of Burwood (0.023), Riccarton South (0.015) and Cashmere West (0.023) all identified by the 1980-2004 analysis, together with St Martins (0.024) to the south of the city.

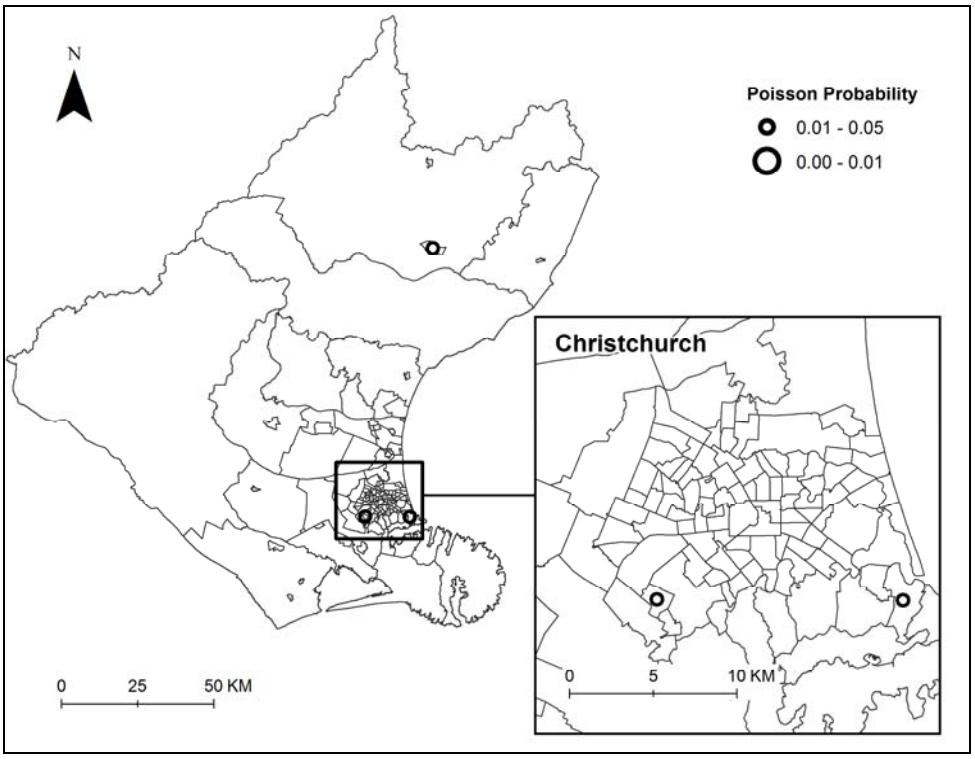

Figure 8.13: Type 1 diabetes Poisson probabilities 1980-1991

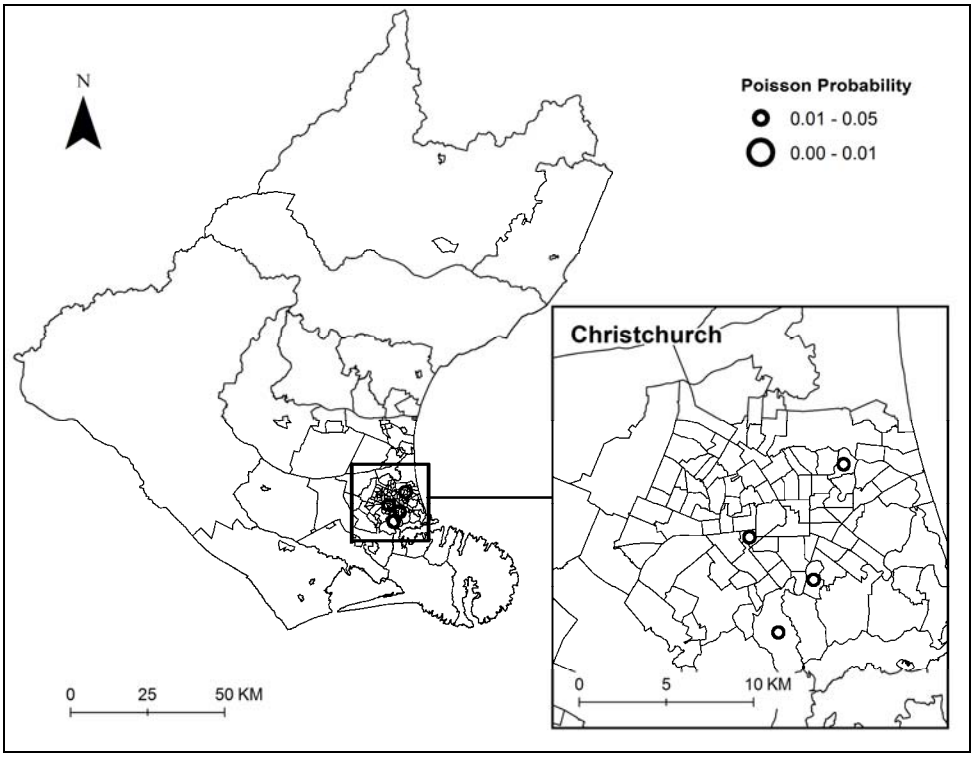

Figure 8.14: Type 1 diabetes Poisson probabilities 1992-2004

8.3.3 Spatial-temporal clustering of type 1 diabetes

Visual analysis of the type 1 diabetes SIRs and Poisson probabilities revealed groups of adjacent areas with similar incidence/probability of this disease. However, these analyses only considered incidence/probabilities over 12, 13 or 25 year time periods, and did not take into account the potential spatial dependence of type 1 diabetes cases. In order to determine whether cases of type 1 diabetes in Canterbury displayed any evidence of spatial clustering over varying time periods, a series of cluster analyses were conducted. The spatial-scan statistic was utilised to test for retrospective space-time clusters using the Poisson probability model (2006). Since it is unclear at what spatial and temporal aggregations type 1 diabetes cases might cluster, a number of runs were conducted. Initially, cases were aggregated to the CAU-level and separate runs were conducted for children of all ages, followed by analyses disaggregated by sex, by age group, and then by sex and age group (Table 8.6). In addition, cases were disaggregated to the smaller meshblock (MB)-level and, separate runs for all ages and for the three age groups, were undertaken. Table 8.6 gives a summary of the results of the CAU-level analyses. Clusters significant at the 95 percent confidence level were found when male and female children were analysed separately, and in males aged 5-9 years.

Table 8.6: Summary of type 1 diabetes cluster analyses results at the CAU-level

Analysis	**P-value of most likely cluster**
All cases:	
0-14 years, males & females	< 0.10
Cases disaggregated by sex:	
0-14 years, males only	< 0.05
0-14 years, females only	< 0.05
Cases disaggregated by age group:	
0-4 years, males & females	< 0.50
5-9 years, males & females	< 0.50
10-14 years, males & females	< 0.50
Cases disaggregated by sex & age group:	
0-4 years, males only	< 1.00
0-4 years, females only	< 0.50
5-9 years, males only	< 0.05
5-9 years, females only	< 1.00
10-14 years, males only	< 1.00
10-14 years, females only	< 1.00

8.3.3.1 All cases

A cluster significant at the 90 percent confidence level was noted when all 337 children were included in the same analysis (Table 8.7 and Figure 8.15). This cluster occurred over a twelve year period (1991-2002) and covered an area with a radius of 26.81km, including adjoining CAUs in southern Christchurch and Banks Peninsula, and two outlying areas; Lincoln and Taitapu. In these areas a total of 60 cases of type 1 diabetes were diagnosed between 1991 and 2002 where only 31.95 cases were expected according to the population at risk, resulting in a relative risk of 2.07. Two secondary clusters were found in the area between southern Rangiora, Waikuku and northern Kaiapoi, and also in north-west Christchurch, but neither was statistically significant.

Table 8.7: Details of clusters of children with type 1 diabetes aged 0-14 years at diagnosis, 1980-2004: CAU-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Obs.	Exp.	Relative Risk
1	S Christchurch & Banks Peninsula	26.81	1991/1/1	2002/12/31	0.0925	60	31.95	2.07
2	S Rangiora, Woodend, Waikuku & N Kaiapoi	6.83	1998/1/1	2004/12/31	0.3846	16	4.43	3.74
3	NW Christchurch	6.89	2002/1/1	2002/12/31	0.9999	3	0.19	15.87

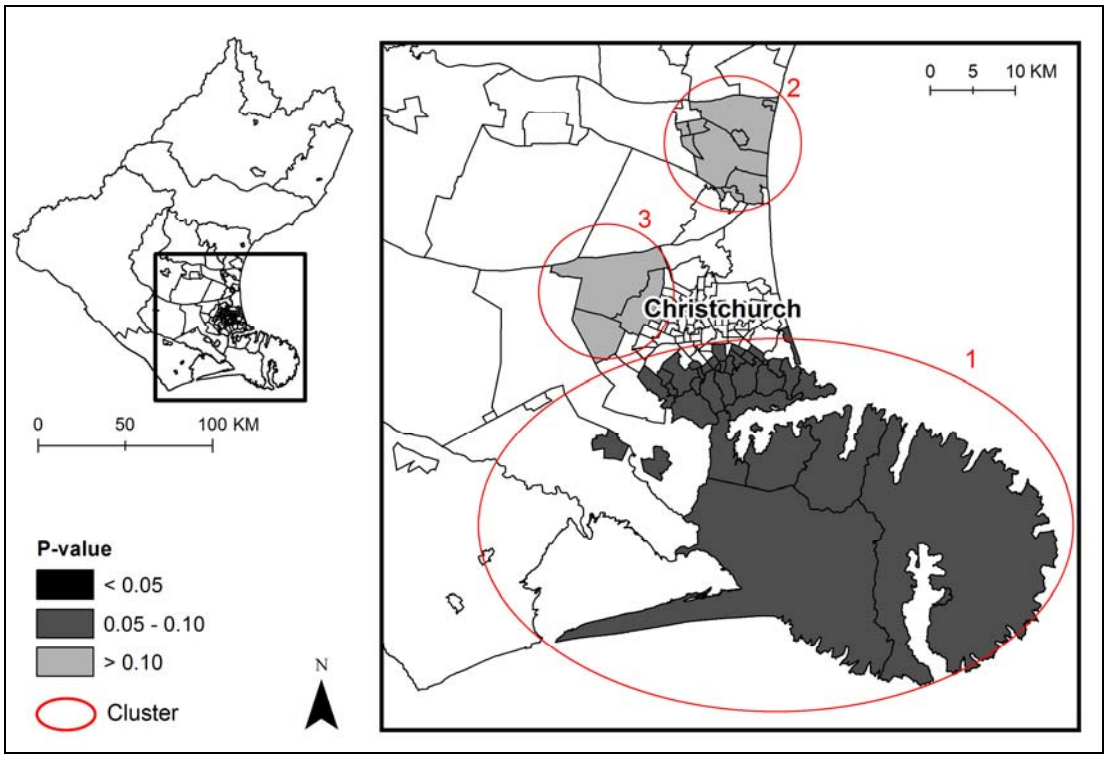

Figure 8.15: Clusters of children with type 1 diabetes aged 0-14 years at diagnosis for the period 1980-2004: CAU-level

8.3.3.2 Clusters by sex

Two significant clusters were found after disaggregating the cases by sex. The most likely cluster of male cases aged 0-14 years at diagnosis occurred between 1998 and 1999 (Table 8.8 and Figure 8.16). The cluster covered the area from the south-east of Rangiora, north-east to Waikuku Beach, and as far south as the north and west fringes of Kaiapoi (total radius of 5.90km). This cluster was based on 6 observed cases where only 0.28 cases were expected according to the age- and sex-specific population at risk. The relative risk for this cluster was thus 22.06 and the cluster's p-value was 0.0207. This area was identified as a non-significant secondary cluster in the analysis which included all 337 children (Table 8.7 and Figure 8.15). Three non-significant secondary clusters of male children with type 1 diabetes were also observed in central and southern Christchurch; West Melton, Springston, Templeton, Prebbleton, Lincoln and Rolleston; and Yaldhurst and Hawthornden (Table 8.8 and Figure 8.16).

The results of the female only analyses highlighted central Christchurch as having a significant excess of type 1 diabetes cases between 1997 and 2004 (Table 8.9 and Figure 8.17). This cluster covered an area with a radius of 8.51km and included CAUs from Marshlands in the northern suburbs, to Middleton in the west, to Lyttelton in the south, and to the coast in the east. Over the

eight year period, 52 new cases of type 1 diabetes were observed in females compared to 26.58 expected cases; a relative risk of 2.37 (p-value = 0.0476). Secondary clusters occurred in the highly rural/remote area of Culverden; rural areas with a moderate urban influence in inland Waimakariri; and the urban areas of Oaklands and Casebrook. However, none were significant. Clusters found in the rural areas of the region for example the inland Waimakariri District Council cluster (Table 8.9), tended to cover large areas due to the small populations living in these areas.

Table 8.8: Details of clusters of male children with type 1 diabetes aged 0-14 years at diagnosis, 1980-2004: CAU-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Obs.	Exp.	Relative Risk
1	SE Rangiora, Waikuku, Woodend & Tuahiwi	5.90	1998/1/1	1999/12/31	0.0207	6	0.28	22.06
2	Central & Southern Christchurch	8.35	1990/1/1	1997/12/31	0.3534	46	24.53	2.22
3	W Melton, Springston, Templeton, Prebbleton, Lincoln & Rolleston	11.00	1999/1/1	2004/12/31	0.3797	10	1.81	5.80
4	Yaldhurst & Hawthornden	2.20	1991/1/1	1991/12/31	0.9994	2	0.04	45.66

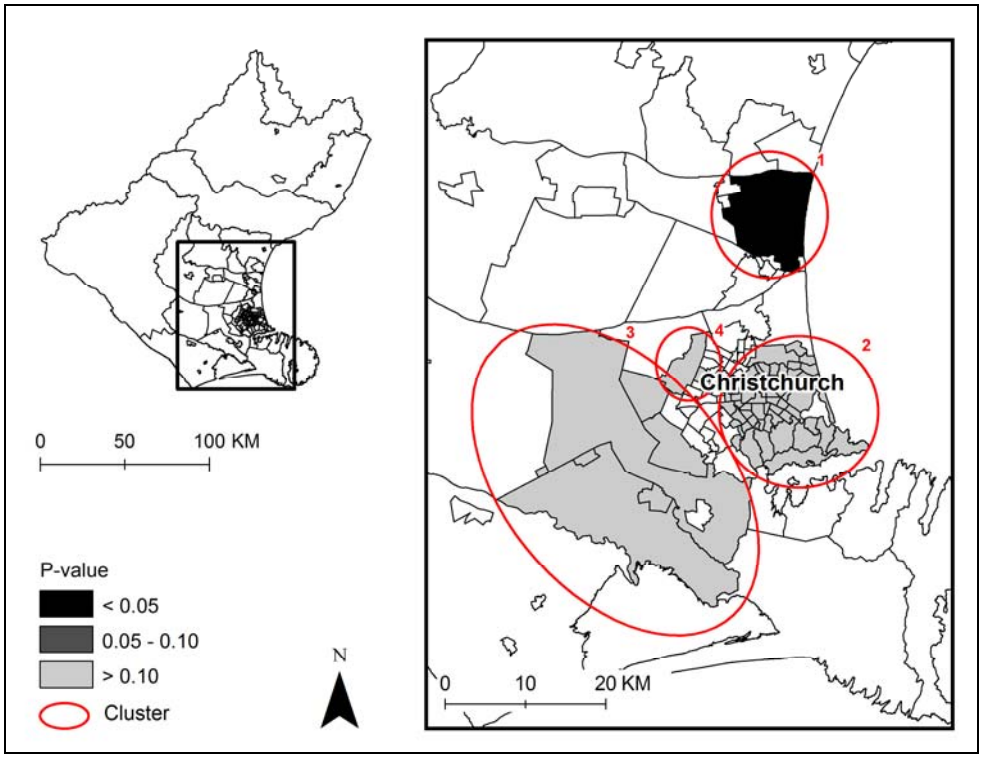

Figure 8.16: Clusters of male children with type 1 diabetes aged 0-14 years at diagnosis for the period 1980-2004: CAU-level

Table 8.9: Details of clusters of female children with type 1 diabetes aged 0-14 years at diagnosis, 1980-2004: CAU-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Obs.	Exp.	Relative Risk
1	Christchurch	8.51	1997/1/1	2004/12/31	0.0476	52	26.58	2.37
2	Culverden	0.00	1986/1/1	1986/12/31	0.6344	2	0.01	161.66
3	Oaklands	0.00	1991/1/1	1992/12/31	0.7243	4	0.22	18.60
4	Inland Waimakariri DC, Malvern & Kirwee	40.43	2004/1/1	2004/12/31	0.9561	5	0.54	9.57
5	Casebrook	0.00	2001/1/1	2001/12/31	0.9992	2	0.04	45.97

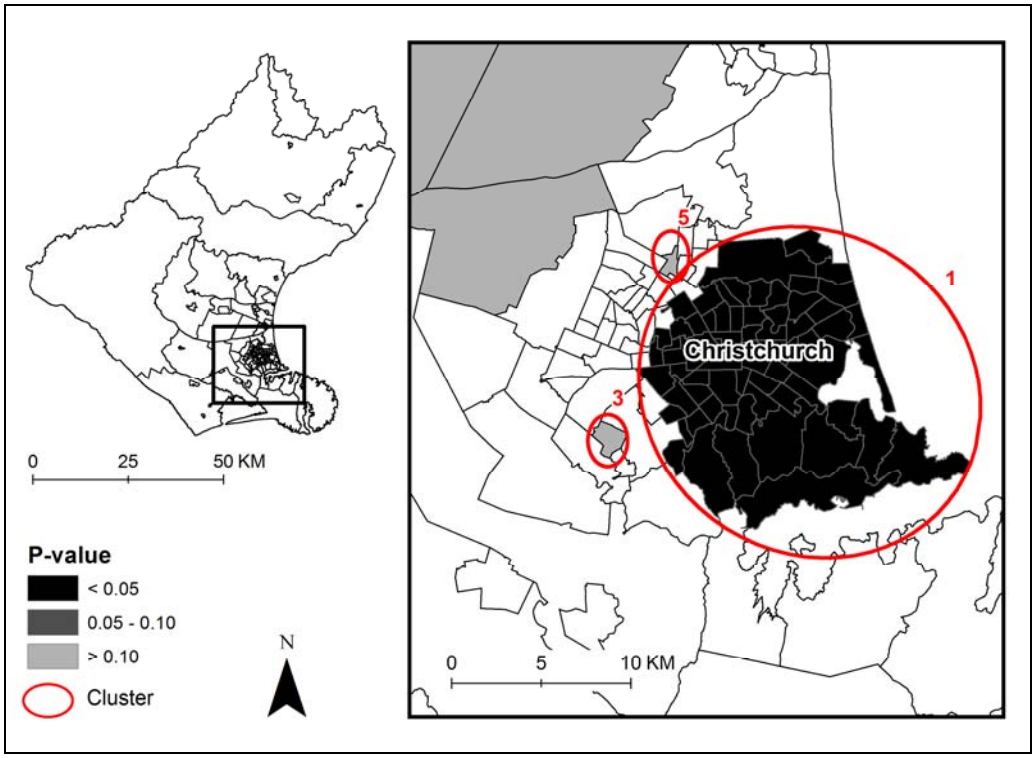

Figure 8.17: Clusters of female children with type 1 diabetes aged 0-14 years at diagnosis for the period 1980-2004: CAU-level

8.3.3.3 Clusters by age group

There were no statistically significant clusters at the CAU-level for children grouped by age group at diagnosis. The p-values of the most likely cluster in all three analyses were between 0.1 and 0.5 (Table 8.6).

8.3.3.4 Clusters by age group and sex

The only significant cluster found when disaggregating the cases by both age group and sex, was for males aged 5-9 years at diagnosis (Table 8.10 and Figure 8.18). This cluster was in a similar geographical location to that noted for the all age male analysis (Table 8.8 and Figure 8.16), although it did not reach as far south, with north and west Kaiapoi no longer included. The cluster covered areas classified as being highly urbanised (Camside), satellite urban communities (East Rangiora, Southbrook, and Woodend) rural areas with a high urban influence (Waikuku Beach and Tuahiwi) and rural areas with a moderate urban influence (Coldstream). The cluster also occurred over the same time period, 1998-1999, during which four males aged 5-9 years were diagnosed with type 1 diabetes when only 0.08 cases were expected. The relative risk for this cluster was much higher (52.14) with a slightly larger, but still statistically significant p-value (0.0407). Non-significant secondary clusters were located in Templeton, Prebbleton, West Melton and Rolleston; central Christchurch; Moncks Bay, Sumner, Lyttelton, Diamond Harbour and Port Levy; and Hornby South, Sockburn and Wigram.

Table 8.10: Details of clusters of male children with type 1 diabetes aged 5-9 years at diagnosis, 1980-2004: CAU-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Obs.	Exp.	Relative Risk
1	SE Rangiora, Waikuku, Woodend & Tuahiwi	5.90	1998/1/1	1999/12/31	0.0407	4	0.08	52.14
2	Templeton, Prebbleton, West Melton & Rolleston	6.06	1999/1/1	2000/12/31	0.8809	3	0.12	26.47
3	Central Christchurch	3.13	2002/1/1	2002/12/31	0.9257	4	0.32	13.58
4	Moncks Bay, Sumner, Lyttelton, Diamond Harbour & Port Levy	11.45	1991/1/1	2001/12/31	0.9702	5	0.62	8.76
5	Hornby South, Sockburn & Wigram	2.08	1998/1/1	1998/12/31	0.9994	2	0.05	38.88

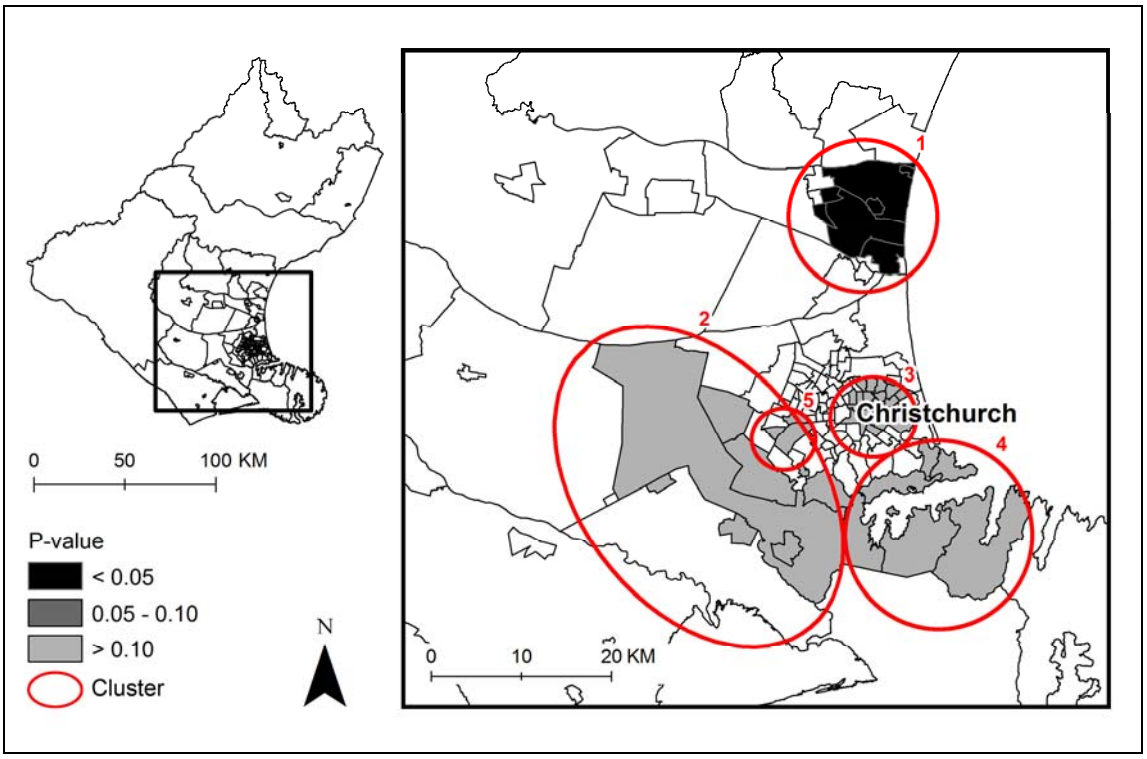

Figure 8.18: Clusters of male children with type 1 diabetes
aged 5-9 years at diagnosis for the period 1980-2004: CAU-level

8.3.3.5 Meshblock analyses

In order to detect potential clusters occurring at a finer geographical scale, analyses were also conducted at the MB-level. Separate analyses for all children, those aged 0-4, 5-9 and 10-14 years, were carried out and no statistically significant clusters were identified (Table 8.11). For example, from the 0-14 year analyses (Table 8.12 and Figure 8.19) a number of small clusters were observed, but all were more likely to have been a chance occurrence rather than a significant find. The most likely cluster (cluster number 1) in Rangiora East was based on two cases where none were expected between 2000 and 2004 with a p-value of 0.5602.

Table 8.11: Summary of type 1 diabetes clusters analyses results at the Meshblock-level

Analysis	**P-value of most likely cluster**
All cases:	
0-14 years, males & females	< 1.00
Cases disaggregated by age group:	
0-4 years, males & females	< 0.50
5-9 years, males & females	< 0.50
10-14 years, males & females	< 0.50

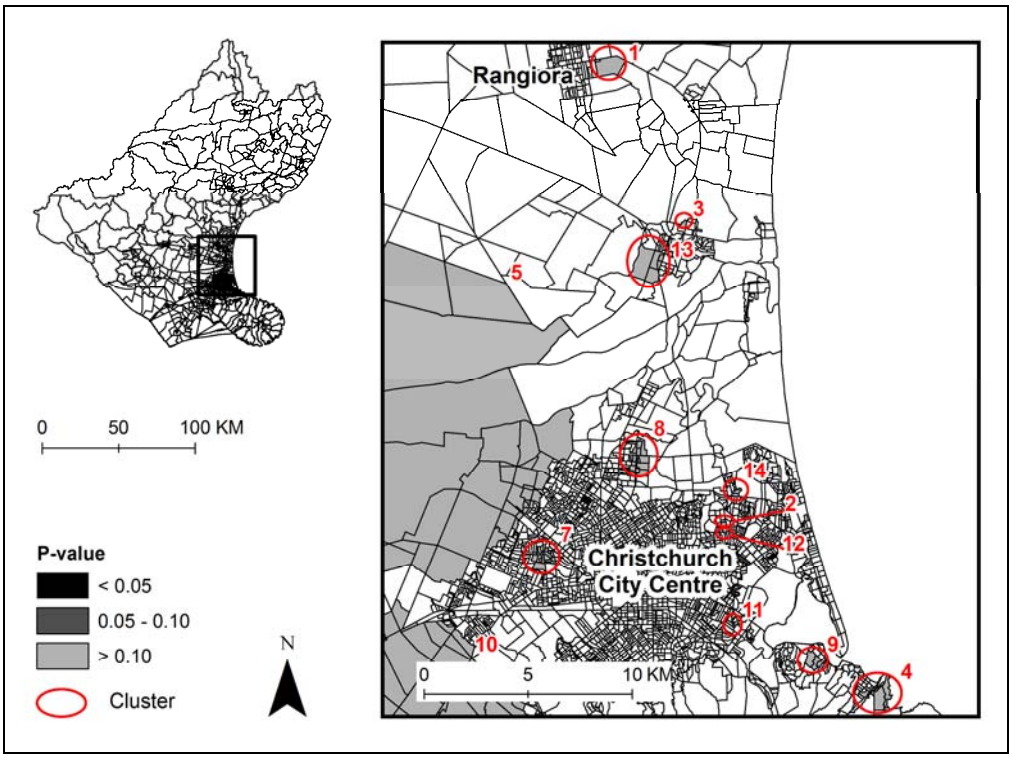

Figure 8.19: Clusters of children with type 1 diabetes aged 0-14 years at diagnosis for the period 1980-2004: Meshblock-level

Table 8.12: Details of clusters of children with type 1 diabetes aged 0-14 years at diagnosis, 1980-2004: Meshblock-level

Cluster Number	Cluster Location	Radius (km)	Start Date	End Date	P-value	Observed	Expected	Relative Risk
1	Rangiora East	0.00	2000/1/1	2004/12/31	0.5602	2	0.00	766.07
2	Burwood	0.00	1995/1/1	1996/12/31	0.7832	2	0.00	531.35
3	Kaiapoi North	0.00	1990/1/1	1990/12/31	0.8432	2	0.00	472.16
4	Sumner	0.81	1997/1/1	1997/12/31	0.8845	3	0.04	78.53
5	NW Christchurch	7.81	1991/1/1	1991/12/31	0.8880	4	0.12	33.11
6	Culverden	0.00	1986/1/1	1986/12/31	0.9404	2	0.01	365.54
7	Ilam & Westburn	0.50	1989/1/1	1991/12/31	0.9753	4	0.14	28.04
8	Redwood North	0.74	1990/1/1	1991/12/31	0.9947	4	0.16	25.00
9	Moncks Bay	0.33	1988/1/1	1992/12/31	0.9994	3	0.06	47.16
10	SW Christchurch	10.12	1999/1/1	2004/12/31	0.9998	13	3.26	4.11
11	Ferrymead	0.00	1998/1/1	1999/12/31	0.9999	2	0.01	171.70
12	Burwood & Dallington	0.30	1999/1/1	2002/12/31	0.9999	3	0.07	41.84
13	Clarkville & Kaiapoi South	0.93	1996/1/1	1996/12/31	0.9999	2	0.01	159.10
14	Marshland	0.00	2001/1/1	2002/12/31	0.9999	2	0.01	158.27

Summary

The geographical analyses described above have consistently highlighted a number of areas as having higher than expected incidence of type 1 diabetes. Within the main urban area of Christchurch, the CAU of Riccarton South was shown to have significantly higher incidence in the latter half of the study period from both the SIR and Poisson probability analyses. In addition, it formed part of a significant cluster of female children noted for the period 1997-2004 at the CAU-level. In Moncks Bay to the south-east of the city there was a higher than average SIR for all periods, significant Poisson probabilities for 1980-2004 and 1980-1991, and the area was also part of the significant cluster of female children in Christchurch during 1997-2004. In addition, the Rangiora, Waikuku Beach and North Kaiapoi triangle observed high SIRs overall and especially in the latter half of the study, and were identified as being part of two significant clusters of males aged 0-14 years and 5-9 years between 1998 and 1999. The only highly rural

area that consistently showed a high incidence was Culverden, as shown by both the SIR and Poisson probability analyses for 1980-1991. This area was also part of two non-significant clusters.

A number of CAUs also had consistently lower than average incidence of type 1 diabetes for the periods 1980-1991 and 1992-2004. Most notably, the CAUs of Ashley Gorge, Okuku, Mairaki and West Eyreton in the centre of the study area witnessed no new cases during the entire study period. The expected number of cases for these CAUs ranged from 0.24 in Mairaki to 1.28 in Ashley Gorge. In addition Port Levy and Okains Bay on Banks Peninsula and Parnassus in Hurunui also had SIRs of zero throughout the study period.

8.4 Type 1 diabetes and population mixing

The results described up to this point have focussed on the descriptive and geographical epidemiology of type 1 diabetes in the Canterbury region and therefore addressed the first aim of the thesis. The final objective of this research was to determine whether there was a relationship between population mixing and type 1 diabetes at the small area-level. To address this aim, a number of methods were applied, ranging from exploratory analyses of SIRs to multivariate generalised linear modelling.

8.4.1 Exploratory analysis

8.4.1.1 Type 1 diabetes by population change and migration change quintile

To explore how type 1 diabetes incidence differed by various population mixing measures, SIRs were initially calculated by population mixing quintiles. CAUs in Canterbury were grouped according to their level of relative population change between 1981 and 2001. Those areas falling into quintile 1 had experienced the highest level of population growth over this period; an increase of between 54.64 and 930.43 percent (Table 5.10, Chapter 5). Quintile 5 areas were those where the population either grew slightly (up to 1.71 percent) or decreased (by up to 35.76 percent). A total of 7 CAUs could not be assigned to a population change or migration change quintile due to small numbers. No cases of type 1 diabetes were noted in these CAUs between 1980 and 2004.

The results show that the highest incidence of type 1 diabetes occurred in quintiles 1 (108.75) and 2 (119.49), where there had been the most growth in population over the study period (Table 8.13). The SIRs for the remaining three quintiles were all below 100, with the smallest SIR occurring in quintile five CAUs (88.41). However, none of the SIRs were statistically significant.

Table 8.13: Type 1 diabetes SIRs, CIs and chi-square values by quintile of population change, 1980-2004

Quintile of Population Change	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
1 High growth	59	54.25	108.75	79.81	137.69	0.42	NO
2	73	61.09	119.49	89.53	149.46	2.32	NO
3	66	70.32	93.86	71.92	115.80	0.26	NO
4	80	84.60	94.56	74.41	114.71	0.25	NO
5 High decline	59	66.74	88.41	67.20	109.62	0.90	NO

The opposite trend was shown for migration change quintiles. The highest incidence was observed in quintiles 4 (109.95) and 5 (137.38); areas where the percentage of total migrants had declined by 8.10 to 68.85 percent (Tables 5.11 Chapter 5, and Table 8.14). Lower SIRs were noted in the first three quintiles, especially in quintiles 1 (89.71) and 3 (84.14). The confidence intervals for all of the SIRs were wide and included 100, suggesting that these groups of areas could all have had either below average, average, or above average type 1 diabetes incidence between 1980 and 2004. The highest SIR of 137.38 (CI = 88.09-186.67) in quintile 5, was the only SIR with a significant chi-square statistic.

Table 8.14: Type 1 diabetes SIRs, CIs and chi-square values by quintile of change in the percentage of total migrants, 1980-2004

Quintile of change in % total migrants	Observed	Expected	SIR	Lower CI	Upper CI	Chi-square	Significant
1 High growth	76	84.71	89.71	70.61	108.82	0.90	NO
2	78	74.60	104.55	80.83	128.28	0.15	NO
3	67	79.63	84.14	65.66	102.62	2.00	NO
4	75	68.21	109.95	83.86	136.05	0.68	NO
5 High decline	41	29.84	137.38	88.09	186.67	4.17	YES

8.4.1.2 Correlation analysis

In addition to considering how type 1 diabetes incidence varied by population mixing categories, simple correlation analyses were conducted. Correlation analyses were carried out in order to determine the strength and directions of bivariate relationships between type 1 diabetes, and each

of the control and population mixing variables. The associations between the population mixing variables and the control variables were also assessed.

The relationships found between type 1 diabetes and the population mixing measures were not particularly strong (Table 8.15). The largest correlation coefficient noted was 0.277 for the variable 'percentage change in total migrants'. This value reveals that as the percentage of total migrants in an area increased, so too did the number of type 1 diabetes cases diagnosed. This relationship was statistically significant at the 99 percent confidence level. A similar, but slightly weaker relationship was noted for the percentage change in child migrants (0.267). Unsurprisingly, the variables representing the total numbers of all-age and child migrants entering an area also showed a weak positive relationship with the type 1 diabetes count. Very weak and non-significant associations were found between the disease and population change, the percentage change in overseas migrants, and also the percentage change in migrant diversity between 1980 and 2004.

Table 8.15: Spearman's rank order correlation coefficients between type 1 diabetes count and population mixing variables, 1980-2004

Population mixing variable	**Correlation Coefficient**
Change in the number of total migrants	0.273**
Change in the number of child migrants	0.197*
Population change	-0.140
Change in the percentage of total migrants	0.277**
Change in the percentage of child migrants	0.267**
Change in the percentage of overseas migrants	0.004
Change in migrant diversity	-0.104

** denotes statistically significant at the 99% confidence level
* denotes statistically significant at the 95% confidence level

In terms of the control variable correlations, the count of type 1 diabetes was positively associated (0.637) with the population of 0-14 year olds, and population density (0.314) at the 1991 census. Both of these relationships were statistically significant at the 99 percent confidence level. No significant linear relationships were noted between type 1 diabetes and the percentage of Europeans, deprivation score and household overcrowding variables (Table 8.16).

Table 8.16: Spearman's rank order correlation coefficients between type 1 diabetes count and control variables, 1980-2004

Control variable	**Correlation Coefficient**
Population 1991	0.637**
% European	-0.067
Population Density	0.314**
Deprivation Score	0.050
% Households < 3 people	0.012
% Households > 6 people	-0.014

** denotes statistically significant at the 99% confidence level
* denotes statistically significant at the 95% confidence level

The population of 0-14 year olds in 1991 was positively associated with four of the population mixing variables; change in the number of total migrants (0.455), change in the number of child migrants (0.275), change in the percentage of total migrants (0.424) and change in the percentage of child migrants (0.371) (Table 8.17). This variable observed a weak negative association with population change (-0.206), and no associations with change in the percentage of overseas migrants or change in migrant diversity. Furthermore, none of these variables were significantly related to any of the control variables for the period 1980-2004. A similar pattern of results was noted between the population mixing measures and the control variables population density, deprivation score and the percentage of households with less than three usual members. The population mixing measures related to the change in the number/percentage of incoming migrants were positively associated with the control variables, whereas population change was inversely associated with these control variables. The opposite trend was visible for the

percentage of Europeans by area; as the percentage of Europeans increased, the change in in-migration variables decreased, but population change increased. The percentage of households with more than six usual members was only significantly related to the change in the percentage of child migrants (-0.180).

Table 8.17: Spearman's rank order correlation coefficients between control and population mixing variables, 1980-2004

	Number of total migrants	Number of child migrants	Population Change	Change % migrants	Change % child migrants	Change % overseas migrants	Change migrant diversity
Population 1991	0.455**	0.275**	-0.206**	0.424**	0.371**	0.046	-0.132
% European	-0.275**	-0.094	0.336**	-0.392**	-0.351**	0.088	0.124
Population Density	0.326**	0.216**	-0.433**	0.469**	0.500**	0.116	-0.145
Deprivation Score	0.178*	0.040	-0.390**	0.405**	0.374**	-0.126	-0.142
% Households < 3 people	0.153	0.091	-0.261**	0.306**	0.346**	-0.004	-0.107
% Households > 6 people	-0.085	-0.054	0.102	-0.124	-0.180*	0.013	0.096

** denotes statistically significant at the 99% confidence level
* denotes statistically significant at the 95% confidence level

8.4.2 Regression modelling

The results reported thus far have been obtained using descriptive and exploratory methods. The relationships between different variables and type 1 diabetes incidence were either inspected alone or in conjunction with the population at risk. No single cause of type 1 diabetes has been identified, and multiple agents could operate synergistically to initiate the onset of the disease (Haverkos, 1997, Leslie and Elliott, 1994). Consequently, it is also necessary to examine several possible explanatory variables at the same time.

Multivariate regression analyses were therefore conducted to simultaneously assess the effects of several explanatory variables on the dependent variable; type 1 diabetes count. Since the majority of the CAUs in Canterbury only observed between 0 and 3 new cases of type 1 diabetes

during the whole study period (25 years), Poisson regression was utilised as the starting model from which to analyse these data.

The modelling was undertaken in six stages (Figure 8.20). The results of steps 1 and 2 were discussed in Chapter 5 (Data and methods), and the results of steps 3 through to 6 are reported in the remainder of this chapter. Univariate model results are presented first (steps 3 and 4) followed by a description of how the multivariate models were formulated (step 5). The final section details the results of the multivariate models (step 6). Analyses were conducted for the study period as a whole (1980-2004), two time periods of 12 and 13 years (1980-1991 and 1992-2004), and four time periods of 6/7 years (1980-1986, 1987-1992, 1993-1998 and 1999-2004) (Table 8.18).

Summary of the modelling strategy:

1. Initially, the appropriate type of probability distribution was chosen to model the data. Examination of the descriptive statistics revealed that the Poisson distribution was an appropriate starting model. Thus, at the outset, Poisson models were fitted to random combinations of variables. These models were then statistically compared with negative binomial and zero-inflated models to determine which model better fitted the data.
2. After the appropriate type of regression model was selected, each explanatory variable was examined to check whether they should be transformed for inclusion in the models.
3. In the first instance, models containing only an intercept were fitted in order to provide a measure of the variation in the dependent variable around its mean.
4. Next all of the explanatory variables were then modelled separately. The age and sex-specific population at the mid-point of the study period was included as an exposure variable in each model and robust standard errors were calculated to account for within area clustering.
5. Using theory guided by the literature and the results of the univariate control models for the 12/13 and 25 year analyses, a base model and alternative multivariate models were constructed (Figure 8.21). Where univariate results for the year group being analysed indicated additional significant control variables, the base and alternative models were altered accordingly.
6. Each population mixing measure was then added separately to each of these models.

Figure 8.20: Summary of regression modelling strategy
(See chapter 5 for more detail)

Table 8.18: Type 1 diabetes analysis time periods

Time period	Number of years	Total cases
1980-2004	25	337
1980-1991	12	111
1992-2004	13	226
1980-1986	7	55
1987-1992	6	72
1993-1998	6	90
1999-2004	6	120

8.4.2.1 Univariate model results

This section details the results of univariate regression models which explored how childhood type 1 diabetes varied by time period, individual- and area- level control variables, and a variety of population mixing measures. Each of the variables was modelled alone, with the age- and sex-specific population at risk included as an exposure variable. Since these data were not significantly over-dispersed or zero-inflated, Poisson regression models were appropriate for all of the models examined. Associations between the type 1 diabetes counts and the explanatory variables were calculated as incidence rate ratios (IRRs).

Change over time

Change in type 1 diabetes incidence over time was firstly assessed. Figures 8.4-8.6 in section 8.2 of this chapter revealed that the incidence of type 1 diabetes in Canterbury children rose considerably between 1980 and 2004. Poisson regression analyses were employed to test whether these increases were statistically significant. Cases of type 1 diabetes were grouped by year group at diagnosis for the following periods: 1980-1991 and 1992-2004 (12/13 years); and 1980-1986, 1987-1992, 1993-1998 and 1999-2004 (6/7 years) (Table 8.18). Year group at diagnosis was then modelled as a categorical explanatory variable.

Results for the 12/13 year analysis revealed a significant increase in type 1 diabetes cases between the two time periods (Table 8.19). The IRR of 1.923 (CI = 1.544-2.395) for the period 1992-2004 asserts that around 92 percent more cases of type 1 diabetes were observed between 1992 and 2004, compared to the period 1980-1991 (base category). Disaggregating these data

into 6/7 year groups displayed a similar trend (Table 8.20). The IRRs for the 6/7 year groups increased in a dose-response relationship over time, with the highest IRR noted in the most recent period (1999-2004). The IRR for 1999-2004 shows that just over twice as many cases of type 1 diabetes were observed during this period, compared to the base category, 1980-1986.

Table 8.19: Results of the type 1 diabetes univariate Poisson regression analyses for year group at diagnosis (12/13 years)

Year Group	IRR	Robust SE	P-value	LCI	UCI
1980-1991 (Base Category)	1.000				
1992-2004	1.923	0.215	0.000	1.544	2.395

Table 8.20: Results of the type 1 diabetes univariate Poisson regression analyses for year group at diagnosis (6/7 years)

Year Group	IRR	Robust SE	P-value	LCI	UCI
1980-1986 (Base Category)	1.000				
1987-1992	1.412	0.260	0.061	0.984	2.026
1993-1998	1.654	0.298	0.005	1.162	2.354
1999-2004	2.083	0.320	0.000	1.541	2.816

Control variables

In the univariate analyses of the control variables, only the results from the 12, 13 and 25 year analyses are presented, as these results were utilised to inform the 6/7 year multivariate analyses. The IRRs reported in the null models represent the mean number of cases of type 1 diabetes diagnosed at the CAU-level in Canterbury for that period (Table 8.21). Between 1980 and 2004 an average of 0.353 cases were diagnosed in each CAU (CI = 0.309-0.403). When the study period was divided in two, the average number of cases was highest in the second half of the study (1992-2004). This finding is compatible with the 12/13 year time period results (Table 8.19). During this time, an average of 0.237 (CI = 0.204-0.275) new cases of childhood diabetes occurred in each CAU.

The categorical age group at diagnosis variable showed a statistically significant association with type 1 diabetes counts in Canterbury between 1980 and 2004 (Table 8.21). The IRRs increased with age, with the lowest IRR noted in the youngest age group (1.000), and the highest in the oldest age group (2.117). Those aged 10-14 years at diagnosis witnessed more than twice the

number of cases of type 1 diabetes compared to those in the base category (0-4 years). Children in the middle age group observed 41 percent more cases than those aged 0-4 years. The same trend of increasing IRRs with age group was noted for 1980-1991 and 1992-2004. The oldest age group had significantly higher incidence of type 1 diabetes compared to the youngest age group in both time periods, with the greatest difference between the two age groups found in the first half of the study (10-14 years IRR = 2.165). There were slightly more female cases than males cases in all time periods, with the largest IRR noted in 1992-2004 (Female IRR = 1.101) but these differences were not statistically significant.

Of the area-level variables, the percentage of Europeans was the only one found to be positively related to type 1 diabetes diagnoses (Table 8.21). Areas with a high percentage of residents of European descent had higher incidence of childhood type 1 diabetes. This association was consistent across the study periods and was statistically significant in 1992-2004 and 1980-2004. The area deprivation score was found to be inversely related to childhood diabetes in Canterbury in both the 12/13 and 25 year time periods. For example the IRR of 0.997 for 1980-2004 suggests that rates of this disease were lower in areas of higher material deprivation. Examining deprivation by quintiles and deciles for the period 1980-2004 showed that category one areas (the most affluent) had an increased risk of type 1 diabetes compared with all of the remaining categories which had IRRs below one. Not all of the categories were statistically significant however and most of the confidence intervals included one. As a result, deprivation was kept as a continuous variable in all subsequent models. The continuous deprivation variable was only significantly related to the count of type 1 diabetes cases in the 25 year analysis. Population density was also shown to be negatively associated with type 1 diabetes incidence for all of the time periods analysed and its IRR of 0.999 was statistically significant between 1980 and 2004. For the periods 1992-2004 and 1980-2004, neither of the household overcrowding variables were significant indicators of type 1 diabetes incidence in Canterbury. However, areas with a high percentage of households with more than 6 usual members had significantly lower incidence of type 1 diabetes between 1980 and 1991 (IRR = 0.864, p-value = 0.050).

In terms of the urban/rural status of areas, there was no trend evident in the IRRs across the urban/rural spectrum (Table 8.21). CAUs classed as satellite urban communities had the highest IRR over the 25 year period. The IRR of 1.692 (CI = 1.248-2.293) revealed that 69 percent more cases of diabetes were observed in these areas compared to the base category; main urban areas. A similar IRR (1.667, CI = 1.076-2.583) was noted for the period 1992-2004. However this category was not significantly associated with childhood diabetes between 1980 and 1991. Areas

categorised as rural with a high urban influence also showed an increased risk of type 1 diabetes relative to main urban areas, but this category was only significant in the 1980-2004 analyses (IRR = 1.441, CI = 1.039-1.999). A reduced risk of childhood diabetes was observed in each of the other urban/rural categories for all of the time periods examined. Significantly lower incidence occurred in independent urban communities and rural areas with a low urban influence as no new cases of childhood type 1 diabetes were diagnosed in these areas during the whole study period. Between 1980 and 2004, significantly lower diabetes incidence was also found in rural areas with a moderate urban influence (IRR = 0.669, CI = 0.462-0.971).

The effects of the control variables on type 1 diabetes incidence in Canterbury were generally consistent across the three time periods analysed. In summary, childhood type 1 diabetes was found to be highest in children aged 10-14 years at diagnosis, in areas with a high percentage of Europeans, in areas with lower population densities and lower levels of household overcrowding, in satellite urban communities and rural areas with a high urban influence, and in the most affluent areas of Canterbury.

Table 8.21: Univariate results of the type 1 diabetes Poisson regression models by 12/13 & 25 year group and control variable

Control variable	1980-1991 Poisson	1992-2004 Poisson	1980-2004 Poisson
None (null model)	0.116* (0.096-0.142)	0.237* (0.204-0.275)	0.353* (0.309-0.403)
Age group category			
0-4 years (base)	1.000	1.000	1.000
5-9 years	1.146 (0.669-1.961)	1.293 (0.948-1.764)	1.411* (1.082-1.842)
10-14 years	2.165* (1.365-3.435)	1.736* (1.233-2.443)	2.117* (1.614-2.776)
Sex category			
Male (base)	1.000	1.000	1.000
Female	1.093 (0.771-1.550)	1.101 (0.842-1.439)	1.088 (0.882-1.342)
% European	1.040 (0.995-1.086)	1.023* (1.004-1.043)	1.029* (1.011-1.047)
Deprivation	0.997 (0.994-1.001)	0.998 (0.996-1.000)	0.997* (0.996-0.999)
Population density	0.999 (0.999-1.000)	0.999 (0.999-1.000)	0.999* (0.999-0.999)
% Households <3 people	1.010 (0.985-1.034)	0.995 (0.975-1.014)	0.996 (0.980-1.012)
% Households >6 people	0.864* (0.746-1.000)	0.956 (0.873-1.047)	0.958 (0.874-1.050)
Urban/rural category (7)			
1-Main urban areas (Base)	1.000	1.000	1.000
2-Satellite urban areas	1.362 (0.753-2.464)	1.667* (1.076-2.583)	1.692* (1.248-2.293)
3-Independent urban comm.	0.000* (0.000-0.000)	0.000* (0.000-0.000)	0.000* (0.000-0.000)
4-Rural high urban influence	1.179 (0.608-2.286)	1.373 (0.969-1.944)	1.441* (1.039-1.999)
5-Rural mod urban influence	0.671 (0.309-1.455)	0.603 (0.277-1.312)	0.669* (0.462-0.971)
6-Rural low urban influence	0.000* (0.000-0.000)	0.000* (0.000-0.000)	0.000* (0.000-0.000)
7-Highly remote/rural	0.929 (0.187-4.604)	0.993 (0.472-2.086)	0.865 (0.487-1.536)

IRRs are reported with 95% confidence intervals in the parentheses.
* denotes a p-value of <0.05.

Static population mixing variables

Static population mixing levels, measured at the beginning of the 12 and 13 year time periods, were inconsistently related to type 1 diabetes incidence over time (Table 8.22). For example, during 1980-1991, type 1 diabetes incidence was higher in areas with a high percentage of total migrants, overseas migrants or overseas visitors in 1981. However, for the period 1992-2004, the IRRs for all of these variables were less than one, suggesting the opposite association for this time period. IRRs for the various population mixing categories also showed higher incidence of type 1 diabetes in areas of high population mixing (category 3 and 4 areas) in 1980-1991, but lower incidence in these areas in the latter half of the study period. However, the percentage of child migrants was positively associated with type 1 diabetes incidence over both time periods, whereas migrant diversity was consistently inversely related to this disease. The CIs for these IRRs all included one. The IRR for the percentage of overseas visitors (1.055, CI = 1.005-1.106) for the period 1980-1991, was the only statistically significant IRR observed in these analyses.

This finding suggests that children living in areas with a high percentage of overseas visitors in 1981 were at increased risk of type 1 diabetes between 1980 and 1991.

Table 8.22: Results of the type 1 diabetes univariate Poisson regression analyses of the static population mixing variables for 12/13 year time periods

Static population mixing variable	1980-1991	1992-2004
Year of PM Measurement	**1981**	**1991**
% of total migrants	1.005 (0.994-1.017)	0.990 (0.972-1.007)
% of child migrants	1.012 (0.966-1.061)	1.009 (0.930-1.095)
% of overseas migrants	1.039 (0.942-1.147)	0.988 (0.903-1.080)
% of overseas visitors	1.055*(1.005-1.106)	0.992 (0.939-1.048)
Migrant diversity score	0.800 (0.555-1.152)	0.913 (0.635-1.313)
Population mixing category:		
1. Low in-migration & low diversity	1.000 (Base)	1.000 (Base)
2. Low in-migration & high diversity	0.889 (0.552-1.433)	0.873 (0.566-1.347)
3. High in-migration & low diversity	1.082 (0.713-1.641)	0.975 (0.732-1.297)
4. High in-migration & high diversity	1.224 (0.529-2.833)	0.879 (0.576-1.337)

IRRs are reported with 95% confidence intervals in the parentheses.
* denotes a p-value of <0.05.

Population mixing change variables

As with the static population mixing measures, the majority of population mixing change variables, were not significant predictors of type 1 diabetes incidence in Canterbury for the periods 1980-1991, 1992-2004 and 1980-2004 (Table 8.23). The only variable with a statistically significant IRR at the 95 percent confidence level was the change in the percentage of total migrants for the period 1980-2004. The IRR was below one (0.997, CI = 0.994-0.999) and thus suggests that there was a small reduction in type 1 diabetes risk in areas where the percentage of migrants had increased the most between 1980 and 2004. A related variable that measured the change in the percentage of child migrants had a similar IRR (0.997, CI = 0.994-1.000) but was only significant at the 90 percent confidence level. Considering these two variables in the 12 and 13 year analyses revealed inconsistent associations over time. Both variables had IRRs below one in the first 12 years of the study, but IRRs greater than one in the final 13 years. A similar trend was noted for the change in the percentage of overseas migrants and overseas visitors. Areas with the greatest increases in population were associated with slightly increased type 1

diabetes incidence in all of the time periods, whereas areas with the greatest increases in migrant diversity had lower incidence in 1980-1991 and 1992-2004, but higher incidence for the study period as a whole (1980-2004).

The categorical population mixing change variable also showed varying directions of association between the different temporal periods analysed. Between 1980 and 1991, type 1 diabetes incidence was higher in areas which had increased in population mixing category (categories 3 and 4), compared to areas which had decreased in population mixing category (base category 1). However, by 1992-2004, areas whose population mixing category had either increased or remained the same (categories 2, 3 and 4), had lower type 1 diabetes IRRs compared to the base category (category 1). During this period, category 3 areas with an IRR of 0.450 (CI = 0.199-1.019) observed approximately 55 percent fewer cases of type 1 diabetes than category 1 areas. This IRR was statistically significant at the 90 percent confidence level. Category 3 and 4 areas also observed lower type 1 diabetes incidence compared to the base category for the period 1980-2004. However, the CIs for these IRRs were wide and included one.

Table 8.23: Results of the type 1 diabetes univariate Poisson regression analyses of the population mixing change variables for 12/13/25 yearly time periods

Population mixing change variable	**1980-1991**	**1992-2004**	**1980-2004**
Population change	1.001 (0.996-1.006)	1.000 (0.995-1.005)	1.001 (0.999-1.001)
Change in % total migrants	0.999 (0.993-1.006)	1.004 (0.995-1.014)	0.997*(0.994-0.999)
Change in % child migrants	0.995 (0.988-1.002)	1.005 (0.998-1.011)	0.997 (0.994-1.000)
Change in % overseas migrants	0.997 (0.993-1.002)	1.002 (0.999-1.003)	1.000 (0.999-1.001)
Change in % overseas visitors	0.999 (0.998-1.000)	1.000 (0.999-1.001)	1.000 (0.999-1.000)
Change in migrant diversity	0.996 (0984-1.009)	0.994 (0.980-1.009)	1.001 (0.998-1.004)
Population mixing change category:			
1. Decrease in either/both categories (base)	1.000	1.000	1.000
2. No change in either category	0.961 (0.620- 1.488)	0.799 (0.603-1.060)	1.007 (0.771-1.315)
3. Increase in diversity category	1.165 (0.569-2.384)	0.450 (0.199-1.019)	0.751 (0.419-1.348)
4. Increase in migration category	1.029 (0.625-1.692)	0.761 (0.503-1.149)	0.870 (0.648-1.167)

IRRs are reported with 95% confidence intervals in the parentheses.
* denotes a p-value of <0.05.

To summarise, the majority of population mixing measures (static and change variables) were not significantly related to type 1 diabetes incidence in Canterbury for the 12/13 or 25 year analyses. The three significant relationships (at the 90-95 percent confidence level) for the population mixing change variables were in a consistent direction. These results suggest that type 1 diabetes incidence was significantly lower in areas which had increased in population mixing during either 1992-2004 or 1980-2004. The only significant association between type 1 diabetes and the static population mixing measures was for areas with a high percentage of overseas visitors in 1981. Incidence of type 1 diabetes was significantly higher in these areas for the period 1980-1991.

The next stage in the analysis involved examining whether the relationships noted between population mixing and childhood type 1 diabetes remained after controlling for potential confounding variables.

8.4.2.2 Formulation of the multivariate models

Using the 12/13 and 25 year control univariate model results (Tables 8.21 and 8.24) and findings from previous studies, a number of multivariate models were constructed (step 5 of the modelling strategy). The main objective of these models was to examine any changes in the relationship between type 1 diabetes and the various population mixing measures, after controlling for other key predictors of the disease. First, a simple base model was formed.

Considering the individual-level control variables, age group at diagnosis was firstly modelled in conjunction with sex for both the 12/13 and 25 year periods. Minimal differences were noted in the IRRs and p-values. The interaction effects between the two variables were also examined, with no significant changes. Consequently, the categorical sex variable was not included in the base model. When modelled alone, age group at diagnosis resulted in a large reduction in deviance (134 units in 1980-2004) and was significantly related to type 1 diabetes incidence in all of the time periods. As a result, age group at diagnosis was used together with the age- and sex- specific population at risk to comprise the base model. Each population mixing measure was firstly added separately to this model. The area-level variables percentage European, deprivation score, population density, the percentage of households with more than six usual members and urban/rural classification (7 categories) were also significantly related to type 1 diabetes in these Canterbury data in the 12/13 or 25 year analyses. Furthermore, similar measures have been identified as important predictors of type 1 diabetes in the international literature. Accordingly,

these variables were then added separately to the base model to determine whether their inclusion altered the directions and strengths of the associations between the population mixing measures and childhood diabetes. The final models are shown in Figure 8.21.

Table 8.24: Comparison of type 1 diabetes univariate control models to the Null model 1980-2004

Model	Deviance	Reduction in Deviance
Null	1,482	-
Age group	1,348	134
Urban/Rural (7 Categories)	1,367	115
% Europeans	1,372	110
Deprivation Score	1,372	110
Population Density	1,374	108
> 6 household members	1,379	103
Sex	1,380	102
< 3 household members	1,380	102
Urban/Rural (2 Categories)	1,380	102

Figure 8.21: Type 1 diabetes regression models

Regression models:

Model 1:	Univariate	Population mixing measure (+ age- + sex-specific population at risk as an exposure variable)
Model 2:	Base model	Age group at diagnosis + population mixing measure (+ age- + sex-specific population at risk as an exposure variable)
Model 3:	Alternative	Base model + % European
Model 4:	Alternative	Base model + deprivation score
Model 5:	Alternative	Base model + population density
Model 6:	Alternative	Base model + urban/rural classification (7 categories)
Model 7:	Alternative	Base model + % households with more than 6 usual members

These regression models were used as a template for all of the time periods analysed (1980-2004; 1980-1991 and 1992-2004; 1980-1986, 1987-1992, 1993-1998 and 1999-2004). Thus the control variables which had a significant effect on type 1 diabetes incidence in the 12/13 or 25 year study periods, were included in the 6/7 year multivariate models even where no significant effect was found in the 6/7 univariate results. These variables were included since they have all been linked with childhood diabetes in the literature (see Chapter 4) and shown by the 12/13 or 25 year analyses to be important in the Canterbury setting. In addition, control variables found to be significant in a 6/7 year univariate analysis (but not identified as being significant in the 12/13 or 25 year analysis), were also tested in the models.

8.4.2.3 Multivariate model results

This section describes the results of the multivariate models and is divided into two main sections. The multivariate results are presented for the static population mixing measures (12/13 and 6/7 year analyses) first, followed by the population mixing change measures (25, 12/13 and 6/7 year analyses).

Static population mixing variables

The inclusion of control variables had little effect on the associations between the static population mixing measures and the count of type 1 diabetes cases for the 12/13 year analyses (Table 8.25). As in the univariate analysis, the IRRs for the percentage of total migrants' variable were greater than one during the period 1980-1991, but less than one in the second half of the study period. A high percentage, of either child or overseas migrants at the beginning of the study period, was generally associated with an increased risk of type 1 diabetes for 1980-1991 and 1992-2004. The migrant diversity variable was the only static population mixing measure to have IRRs consistently below one in all of the multivariate models, in both time periods. However the IRR values were not statistically significant. The addition of potential confounding variables did not considerably alter the IRRs noted for the population mixing categories for either period. Furthermore, only one significant association was observed. Between 1980 and 1991, the percentage of overseas visitors variable had a significantly raised IRR in all of the models tested (Figure 8.21), with the exception of model 7. This finding was consistent with the univariate model for this variable, and suggests that type 1 diabetes incidence (1980-1991) was higher in areas with a high percentage of overseas visitors in 1981 (IRR for model 3 = 1.055, CI = 1.008-1.104). However, by 1992-2004, the IRRs for this variable were consistently below one, with high p-values and CIs that included one (IRR for model 3 = 0.990, CI = 0.928-1.057).

Table 8.25: Results of the type 1 diabetes multivariate Poisson regression models by 12/13 year group and static population mixing variable

Static population mixing variable	1980-1991 Poisson	1992-2004 Poisson
YEAR OF PM MEASUREMENT	**1981**	**1991**
% of total migrants	1.004 (0.993-1.015) SIG: NONE	0.996 (0.979-1.014) SIG: NONE
% of child migrants	1.004 (0.956-1.054) SIG: NONE	1.003 (0.923-1.089) SIG: NONE
% of overseas migrants	1.062 (0.957-1.178) SIG: NONE	1.019 (0.936-1.109) SIG: NONE
% of overseas visitors	1.055*(1.008-1.104) SIG: M2,3,4,5,6 (+)	0.990 (0.928-1.057) SIG: NONE
Migrant diversity score	0.852 (0.584-1.243) SIG: NONE	0.973 (0.678-1.398) SIG: NONE
Population mixing category:		
1. Low in-migration & low diversity (Base)	1.000	1.000
2. Low in-migration & high diversity	0.893 (0.568-1.404) SIG: NONE	0.884 (0.584-1.339) SIG: NONE
3. High in-migration & low diversity	1.033 (0.676-1.577) SIG: NONE	1.017 (0.764-1.352) SIG: NONE
4. High in-migration & high diversity	1.484 (0.592-3.720) SIG: NONE	0.962 (0.649-1.425) SIG: NONE

IRRs are reported for population mixing measures in model 3 (age & % European). * denotes a p-value of <0.05. 95% confidence intervals are reported in the parentheses. SIG denotes significant models, and the direction of the PM association in these models (+ or -) is noted in the parentheses.

Disaggregating the diabetes counts by 6/7 year time periods also revealed a significant association between the percentage of overseas visitors and type 1 diabetes incidence, prior to 1992 (Table 8.26). After 1992, the IRRs for this variable were consistently below one in the 1993-1998 models and consistently above one in the 1999-2004 models. The CIs in these models all included one. Furthermore this variable was only significant in the base model (model 2) prior to 1992. Thus similar trends were noted, but the IRRs for this variable were statistically significant in fewer models compared to the 12 year multivariate analysis (1980-1991).

The 6/7 year multivariate analyses also identified two additional static population mixing variables as being significant predictors of type 1 diabetes incidence in Canterbury. Areas with a high percentage of overseas migrants in 1981 had significantly higher type 1 diabetes incidence (IRR = 1.499, CI = 1.157-1.943) between 1980 and 1986 when modelled with age group at

diagnosis and population density (model 5). After this time, the IRRs for this variable fluctuated around one, and were not statistically significant. In addition, areas classed as having high in-migration and high migrant diversity (category 4 areas) had significantly higher childhood diabetes during this time period compared to areas with low in-migration and low migrant diversity (base category 1 areas). For example, when modelled with age group and the percentage of Europeans variable (model 3), the IRR of 3.438 (CI = 1.281-9.225) suggests that over three times as many cases were diagnosed in the high population mixing areas (category 4) compared to the low population mixing areas (base category 1). However, in the following two time periods the IRRs for this category were mostly below one, and the CIs for all of the multivariate models after 1986, included one. Significantly lower type 1 diabetes was observed between 1987 and 1992 in areas categorised as having low in-migration and high migrant diversity in 1986 compared to base category areas (IRR for model 3 = 0.172, CI = 0.042-0.706). This relationship was consistent in all of the models tested but was not significant in other time periods. In general the IRRs of the various population mixing categories did not show consistent patterns over time. The most consistent IRRs were noted for the percentage of total migrants' variable. Areas with a high percentage of total migrants were found to have a non-significant reduced risk of type 1 diabetes during every 6/7 year time period. The results of the child migrants' models generally contradicted this finding, with IRRs greater than one in every period except 1999-2004. However, the CIs all included one.

In summary, the results for the significant static population mixing variables in the 6/7 year analyses were in a consistent direction. There was higher type 1 diabetes incidence in areas where population mixing was higher at the beginning, or just before, the diagnosis period. No significant associations were observed after 1992; this finding is in agreement with the 12 and 13 year multivariate analyses.

Table 8.26: Results of the type 1 diabetes multivariate Poisson and zero-inflated Poisson regression models by 6/7 year group and static population mixing variable

Static population mixing variable	1980-1986 (ZIP)	1987-1992 (ZIP)	1993-1998 (Poisson)	1999-2004 (Poisson)
% of total migrants	0.996 (0.979-1.014) SIG: NONE	0.998 (0.971-1.025) SIG: NONE	0.991 (0.960-1.022) SIG: NONE	0.995 (0.967-1.023) SIG: NONE
% of child migrants	1.003 (0.937-1.074) SIG: NONE	1.036 (0.864-1.243) SIG: NONE	1.006 (0.874-1.157) SIG: NONE	0.940 (0.836-1.056) SIG: NONE
% of overseas migrants	1.215 (0.967-1.528) SIG: M5 (+)	0.999 (0.836-1.195) SIG: NONE	1.048 (0.918-1.197) SIG: NONE	0.979 (0.914-1.048) SIG: NONE
% of overseas visitors	1.013 (0.870-1.179) SIG: M2 (+)	1.000 (0.779-1.282) SIG: M2 (+)	0.953 (0.810-1.122) SIG: NONE	1.011 (0.937-1.091) SIG: NONE
One year mobility %	No Data	0.994 (0.952-1.039) SIG: NONE	0.993 (0.951-1.038) SIG: NONE	1.000 (0.973-1.028) SIG: NONE
Migrant diversity score	0.839 (0.465-1.514) SIG: NONE	0.707 (0.354-1.412) SIG: NONE	1.052 (0.592-1.869) SIG: NONE	1.051 (0.627-1.762) SIG: NONE
Population mixing category:				
1. Low in-migration & low diversity	1.000	1.000	1.000	1.000
2. Low in-migration & high diversity	0.842 (0.439-1.615) SIG: NONE	0.172*(0.042-0.706) SIG: ALL (-)	1.204 (0.608-2.384) SIG: NONE	0.865 (0.394-1.899) SIG: NONE
3. High in-migration & low diversity	0.811 (0.403-1.628) SIG: NONE	0.659 (0.354-1.226) SIG: NONE	1.071 (0.643-1.784) SIG: NONE	0.921 (0.606-1.401) SIG: NONE
4. High in-migration & high diversity	3.438*(1.281-9.225) SIG: M3,5 (+)	0.564 (0.207-1.534) SIG: NONE	0.927 (0.479-1.794) SIG: NONE	1.010 (0.604-1.690) SIG: NONE

IRRs are reported for population mixing measures in model 3 (age & % European).* denotes a p-value of <0.05. 95% confidence intervals are reported in the parentheses. SIG denotes significant models, and the direction of the PM association in these models (+ or -) is noted in the parentheses.

Population mixing change variables

The results of the 12/13 and 25 year multivariate population mixing change models are given in Table 8.27. In general, the findings were similar to those shown in the univariate models (Table 8.23); the majority of population mixing change variables were not significantly associated with childhood type 1 diabetes.

As in the univariate analyses, there were no significant associations between the population mixing change variables and type 1 diabetes during the first 12 years of the study (1980-1991). Furthermore, the directions of the relationships tended to change over time. For example, the IRRs for the change in the percentage of child migrants and the change in the percentage of overseas migrants were below one in every model between 1980 and 1991, but were greater than one in all of the models for the period 1992-2004. The inverse association between type 1 diabetes and the change in migrant diversity was the only association which remained in the same direction in all of the multivariate models by 1992-2004 (IRR for model 3 = 0.992, CI = 0.979-1.006). Although, for the study period as a whole (1980-2004), the IRRs for this variable fluctuated above and below one, depending on the control variables included. In addition, none of these IRRs were statistically significant. The relationships between type 1 diabetes and population change and the change in the percentage of overseas visitors, were not consistent over time and none were statistically significant.

Contrary to the findings for 1980-1991, the IRRs for the population mixing change categories 2, 3 and 4 for the period 1992-2004 were below one for all of the models tested. This finding implies that childhood diabetes was lower in areas which witnessed either no change or an increase in population mixing category, compared to areas which decreased in population mixing category during this period (category 1). The IRRs for category 3 areas were consistently below one and were all significant at the 95 percent confidence level. When age group and urban/rural category were controlled for (model 6), the IRR was 0.486 (CI = 0.254-0.929) and asserts that around 51 percent fewer cases of type 1 diabetes were observed in category 3 areas which had increased in migrant diversity, compared to the base category. This category was almost significant in the univariate regression analyses during 1992-2004. For the study period as a whole, the IRRs for this category remained below one in every model, but were not statistically significant (IRR for model 3 = 0.750, CI = 0.435-1.290).

Also consistent with the univariate analyses, were the inverse associations noted between the change in the percentage of total or child migrants and childhood diabetes, for the period 1980-2004. The IRRs for these two variables remained below one in all of the models, asserting that there was a reduced risk of type 1 diabetes in areas which had increased the most in the percentage of total or child migrants during this time. The IRRs for the change in total migrants variable were statistically significant at the 90 percent confidence level in models 2 (base) and 6 (urban/rural), and for the change in child migrants variable in model 1 (univariate) only. Addition of the other control variables increased the p-values. These variables were not found to be consistently related to type 1 diabetes over the 12 and 13 year time frames.

Thus the significant findings for the 13 year (p-value <0.05) and 25 year (p-value <0.10) multivariate population mixing change analyses were in a compatible direction. These findings suggest that there was a reduced risk of childhood type 1 diabetes in areas which had increased in population mixing between 1992-2004 and 1980-2004.

Table 8.27: Results of the type 1 diabetes multivariate Poisson regression models by year group and population mixing change variable

Population mixing change variable	1980-1991 Poisson	1992-2004 Poisson	1980-2004 Poisson
Population change	1.000 (0.995-1.005) SIG: NONE	0.997 (0.991-1.003) SIG: NONE	1.000 (0.999-1.000) SIG: NONE
Change in % total migrants	1.000 (0.994-1.006) SIG: NONE	1.007 (0.997-1.016) SIG: NONE	0.998 (0.995-1.001) SIG: NONE
Change in % child migrants	0.996 (0.989-1.002) SIG: NONE	1.004 (0.998-1.010) SIG: NONE	0.998 (0.995-1.001) SIG: NONE
Change in % overseas migrants	0.997 (0.993-1.001) SIG: NONE	1.001 (0.999-1.003) SIG: NONE	0.999 (0.998-1.001) SIG: NONE
Change in % overseas visitors	0.999 (0.998-1.000) SIG: NONE	1.000 (0.999-1.001) SIG: NONE	1.000 (0.999-1.000) SIG: NONE
Change in migrant diversity	0.995 (0.984-1.007) SIG: NONE	0.992 (0.979-1.006) SIG: NONE	0.999 (0.996-1.003) SIG: NONE
Population mixing change category:			
1. Decrease in either/both categories (base)	1.000	1.000	1.000
2. No change in either category	1.007 (0.661-1.532) SIG: NONE	0.817 (0.630-1.061) SIG: NONE	0.977 (0.765-1.249) SIG: NONE
3. Increase in diversity category	1.223 (0.658-2.274) SIG: NONE	0.427*(0.194-0.940) SIG: M3,4,5,6 (-)	0.750 (0.435-1.290) SIG: NONE
4. Increase in migration category	1.081 (0.661-1.768) SIG: NONE	0.771 (0.524-1.134) SIG: NONE	0.912 (0.690-1.205) SIG: NONE

IRRs are reported for population mixing measures in model 3 (age & % European). * denotes a p-value of <0.05. 95% confidence intervals are reported in the parentheses. SIG denotes significant models, and the direction of the PM association in these models (+ or -) is noted in the parentheses.

Prior to 1986, population change was inversely related to type 1 diabetes count, as was the change in the percentage of child migrants (Table 8.30). These relationships were only statistically significant (at the 90-95 percent confidence level) when age group at diagnosis and population density were controlled for in the models (model 5). In addition, the CIs for the change in the percentage of child migrants' variable included one in every model. Between 1987 and 1992, population change became positively associated with childhood diabetes when age group and the percentage of Europeans (models 2 and 3) were controlled for (IRR for model 3 = 1.019, CI = 1.002-1.036). Thus in contrast to the earlier period, areas which had grown in population had significantly higher type 1 diabetes incidence between 1987 and 1992. Furthermore, between 1993 and 1998, an increase in the percentage of overseas visitors by area

was associated with significantly higher childhood diabetes incidence over this period. This finding was consistent in all of the models tested. However, the effect size for this variable was small (IRR for model 3 = 1.001, CI = 1.000-1.001).

In the most recent period (1999-2004), an increased risk of childhood diabetes was observed in areas which had increased the most in the percentage of total or child migrants. The change in the percentage of total migrants' variable was only significant when modelled together with age group, sex and the percentage of Europeans (model 3). The change in the percentage of child migrants' variable was significant in all of the models with the exception of models 6 and 7. Therefore areas which had increased the most in the percentage of total or child migrants had higher incidence of childhood diabetes between 1999 and 2004. These results contradict the finding for the whole study period (1980-2004) which showed a decreased risk of type 1 diabetes to be associated with these areas. An increase in the percentage of overseas migrants was also associated with higher type 1 diabetes incidence for the period 1999-2004. This finding was significant at the 90 percent confidence level in models 1, 2, 3 and 4 (IRR for model 3 = 1.002, CI = 0.999-1.004). Furthermore, type 1 diabetes was significantly higher in areas where the percentage of people living there for less than one year had increased during this time. For example, when modelled with age group, sex and the percentage of Europeans (model 3), the IRR for this variable was 1.013 (CI = 1.005-1.020).

In summary, prior to 1987, lower type 1 diabetes incidence was observed in areas where population mixing (population change or change in the percentage of child migrants) had increased the most. However, this finding was only significant in one model (model 5) and the CIs for the child migrants' variable all included one. After 1987, higher type 1 diabetes incidence was associated with areas which had increased in population mixing. This was especially the case in the most recent time period (1999-2004) where a significantly increased risk of type 1 diabetes was noted for increases in three different population mixing change variables (change in the percentage of total migrants, change in the percentage of child migrants, and change in the one year mobility percentage). These findings generally contradict the 12/13 and 25 year multivariate analyses results.

Table 8.28: Results of the type 1 diabetes Poisson and zero-inflated Poisson multivariate regression models by 6/7 year group and population mixing change variable

Population mixing change variable	1980-1986 (ZIP)	1987-1992 (ZIP)	1993-1998 (Poisson)	1999-2004 (Poisson)
Population change	0.970 (0.935-1.006) SIG: M5 (-)	1.019*(1.002-1.036) SIG: M2,3 (+)	0.987 (0.965-1.010) SIG: NONE	0.999 (0.984-1.014) SIG: NONE
Change in % total migrants	0.999 (0.989-1.009) SIG: NONE	1.008 (0.990-1.026) SIG: NONE	0.996 (0.973-1.019) SIG: NONE	1.017*(1.000-1.035) SIG: M3(+)
Change in % child migrants	0.994 (0.983-1.004) SIG: NONE	1.000 (0.988-1.012) SIG: NONE	1.003 (0.991-1.015) SIG: NONE	1.010*(1.000-1.019) SIG: M2,3,4,5 (+)
Change in % overseas migrants	0.999 (0.994-1.004) SIG: NONE	1.000 (0.994-1.007) SIG: NONE	0.998 (0.993-1.002) SIG: NONE	1.002 (0.999-1.004) SIG: NONE
Change in % overseas visitors	0.999 (0.997-1.001) SIG: NONE	1.000 (0.999-1.002) SIG: NONE	1.001*(1.000-1.001) SIG: ALL (+)	0.999 (0.998-1.000) SIG: NONE
Change in 1 year mobility	No Data	1.005 (0.989-1.021) SIG: NONE	1.005 (0.995-1.015) SIG: NONE	1.013*(1.005-1.020) SIG: ALL (+)
Change in migrant diversity	0.998 (0.975-1.021) SIG: NONE	1.013 (0.986-1.042) SIG: NONE	0.990 (0.962-1.019) SIG: NONE	0.999 (0.979-1.021) SIG: NONE
Change in population mixing category :				
1. Decrease in either/both categories (base)	1.000	1.000	1.000	1.000
2. No change in either category	1.780 (0.856-3.701) SIG: NONE	1.656 (0.391-7.012) SIG: NONE	1.001 (0.480-2.089) SIG: NONE	1.089 (0.674-1.759) SIG: NONE
3. Increase in diversity category	1.570 (0.280-8.790) SIG: NONE	1.678 (0.331-8.487) SIG: NONE	0.305 (0.038-2.438) SIG: NONE	0.864 (0.365-2.046) SIG: NONE
4. Increase in migration category	1.703 (0.532-5.446) SIG: NONE	0.847 (0.166-4.317) SIG: NONE	0.676 (0.272-1.681) SIG: NONE	0.841 (0.384-1.839) SIG: NONE

IRRs are reported for population mixing measures in model 3 (age & % European. Sex is also included in 1999-2004 models).* denotes a p-value of <0.05. 95% confidence intervals are reported in the parentheses. SIG denotes significant models, and the direction of the PM association in these models (+ or -) is noted in the parentheses.

8.5 Conclusion

This chapter has revealed several important themes in the geographical epidemiology of childhood type 1 diabetes in Canterbury, 1980-2004. Firstly, type 1 diabetes has been shown to vary by a number of individual-level characteristics, especially age at diagnosis and ethnicity. Secondly, these analyses have revealed a considerable increase in the incidence of childhood type 1 diabetes over the study period, and extensive variations in incidence by geographical location and area type. For example, the incidence of this disease was found to be significantly raised in rural areas and satellite communities with a high urban influence, and also in the most affluent areas of the region. Thirdly, this chapter has found population mixing to be significantly associated with childhood type 1 diabetes in the first ever analysis of this association in a New Zealand setting.

Static population mixing levels, measured at the beginning of the analysis periods, were positively associated with childhood type 1 diabetes in both the 12/13 and the 6/7 year analyses. However, these associations were only noted using three of the seven population mixing measures, and were only significant in the first half of the study period. A wider range of population mixing change variables were found to be significantly associated with type 1 diabetes incidence in the Canterbury region. However, the general directions of association changed depending on the time scale examined. Over longer periods of time (12/13/25 years), the significant relationships between type 1 diabetes and population mixing suggested a decreased risk of this disease in areas where population mixing had increased the most. Over shorter time intervals (6/7 years) however, and especially after 1986, the majority of the associations revealed the opposite trend. Increased risk of type 1 diabetes was associated with areas where population mixing had increased the most. The remainder of the thesis addresses possible explanations for these findings, and those of the ALL analyses.

Chapter 9: Discussion

9.1 Introduction

The precise causes of childhood ALL and type 1 diabetes remain unknown (Daneman, 2005, Terracini and Maule, 2007). These diseases have been shown to have many epidemiological features in common (Feltbower et al., 2005, Feltbower et al., 2004, Staines, 1996), and infections, introduced through population mixing, have been implicated in the aetiology of both diseases (Kinlen, 1988, Parslow et al., 2001). However, the international findings are mixed and previous studies suffer from a number of theoretical and methodological limitations. Furthermore, the issue of population mixing and health has received very little attention in the Pacific Region. This research addressed some of these insufficiencies by examining the role of population mixing in the geographical epidemiology and aetiologies of ALL and type 1 diabetes in New Zealand.

This chapter explores the eight key findings of this research and is divided into four sections. The first two sections examine the main themes associated with the incidence of ALL across New Zealand and type 1 diabetes in the Canterbury region. The evidence for each key finding is assessed and possible explanations are given. The third section of the chapter makes some general comparisons between the prominent ALL and type 1 diabetes findings, and the final section critically assesses the research, including the data and methods employed.

9.2 ALL key findings

Four key themes regarding the geographical epidemiology, and potentially aetiology of childhood ALL in New Zealand can be identified. First, incidence of the disease increased by approximately 20 percent during the 25 year study period (1980-2004). Second, ALL incidence was generally higher in the urban areas of New Zealand throughout the study period. Third, the more affluent areas of New Zealand tended to have higher incidence of childhood ALL compared to the most deprived areas of the country. Finally, increases in small area population mixing were associated with a significantly elevated risk of ALL for a number of time periods. This finding was consistent over time and was most pronounced for short time periods (6/7 years).

9.2.1 Increase over time

There was an overall increase in the incidence of ALL in New Zealand children between 1980 and 2004. This finding is consistent with the results of similar studies in Europe (Shah and Coleman, 2007, Steliarova-Foucher et al., 2005, Steliarova-Foucher et al., 2004), America (Xie et al., 2003) and Australia (Milne et al., 2007). However, studies in France (Desandes et al., 2004) and the Nordic countries (Hjalgrim et al., 2003) have reported more stable rates. Previous analysis of ALL time trends in New Zealand showed a significant increase in age-standardised incidence rates between 1968 and 1990 (Dockerty et al., 1996b). The results of this research suggest a continued increase in ALL incidence in New Zealand children over time.

One possible explanation for the temporal increases in childhood leukaemia is an improvement in cancer registrations over time. The rare nature of childhood ALL means that small changes in the classification or recording of cases can have a substantial effect on time trends (Adamson et al., 2005). There is currently much debate in the international literature as to whether increases noted (especially over long time periods) are simply an artefact of improved data collection (Adamson et al., 2005, Shah and Coleman, 2007, Steliarova-Foucher et al., 2005). However, while under-ascertainment of cases may have affected incidence prior to the 1970s (Shah and Coleman, 2007), it is unlikely to account for more recent time trends. In this research, ALL cases were obtained from the New Zealand Cancer Registry for which case ascertainment is thought to be virtually complete. For the period 1990-1993, the completeness of ascertainment of childhood leukaemia cases was estimated to be 98.6 percent (Dockerty et al., 1997). Moreover, histological classification of the different types of leukaemia in New Zealand is considered to have been relatively accurate since the late 1960s (Dockerty et al., 1996b). Thus, changes in the diagnosis of the main sub-types of leukaemia are unlikely to explain the increases in ALL noted after 1980.

Since genetics have been identified as a risk factor in the aetiology of childhood ALL (Heath, 1996), it is plausible that an increase in the genetically susceptibly population could be partly responsible for the increases in incidence. However, a strong inherited predisposition is thought to be relevant to a small minority of cases (approximately 5 percent) (Greaves, 2006), and thus temporal changes in environmental risk factors are likely to play a key role. A number of candidate environmental causes of childhood ALL have been examined in the international literature, many of which may have varied over time. New Zealand has no nuclear facilities and fallout from nuclear weapons testing has been decreasing since 1965 (Matthews, 1993). Furthermore, natural sources of background radiation, such as radon, are comparatively low in

New Zealand (National Radiation Laboratory, 1998, Robertson et al., 1988) and are unlikely to have increased substantially over time. It has been suggested that an increase in exposure to electromagnetic fields (EMFs) could be involved in previous increases in ALL. However, the average annual household electricity consumption in New Zealand remained relatively stable between 1977 and 1993 while childhood ALL increased steadily during the same period (Dockerty et al., 1996b). Moreover, causal links between EMFs and ALL have not yet been proven (Dockerty et al., 1998, Dockerty et al., 1999a, Dockerty et al., 1999b, Greaves, 2006). A further possible environmental cause of ALL is exposure to benzene in road traffic emissions (Belson et al., 2007). Recent investigations of benzene concentrations in the New Zealand urban environment suggest an increase over time (Clarkson et al., 1996). However, further work would be necessary to confirm these findings, and to assess what impact any increases may have had on childhood ALL rates.

In most developed countries, increases in ALL also parallel decreases in the incidence of many infectious diseases. For example, a dramatic reduction in measles, mumps, rheumatic fever, tuberculosis and hepatitis A has occurred between the mid 1960s and 2000 in the USA (Bach, 2002). During the period 1973-1998, a statistically significant increase in ALL occurrence was also observed (Xie et al., 2003). The decrease in common infections has occurred as a result of the introduction of vaccinations, improved hygiene (for example the use of disposable nappies and food industry products in baby care) and better socioeconomic conditions (Bach, 2002, Kolb and Elliot, 1994). Reduced exposure to infections in early life is thought to be a key factor in the development of childhood leukaemia (Greaves, 2006, Kinlen, 2004a). Data on common infections in New Zealand for the period 1996-2006, suggest similar decreases (Population and Environmental Health Group, 2006). However, longer term trends would need to be examined to assess this potential association further.

Even if the number of common infections has not decreased appreciably over time, children's exposure to them may have. For example, a temporal decrease in children playing outside within their local neighbourhoods (due to a rising fear of crime for instance) would limit their early infectious exposure from contact with other children. In the UK there is evidence to suggest that children's independent use of public space has decreased since the 1970s (O'Brien et al., 2000). This factor coupled with smaller household sizes could mean that children are not being exposed to as many infections in early life. A steady decrease in the average household size has been noted in New Zealand in recent decades. In 1981 there were on average 3.0 people per household, compared to 2.6 people per household in 2001 (Statistics New Zealand, 2004).

Similar to early infectious exposure, breast-feeding has also been associated with having a protective effect against the development of childhood ALL (Altinkaynak et al., 2006, Kwan et al., 2004, UKCCS Investigators, 2001). At the national-level, the prevalence of breast-feeding increased from approximately 81 to 87 percent between 1980 and 1987. The rate dipped in 1990, increased again to around 87 percent in 1994, then decreased steadily to about 78 percent by 2001 (Ministry of Health, 2002). These patterns do not appear to mirror the peaks and troughs noted in ALL incidence (Figure 7.4 Chapter 7), and the decreases in breast-feeding since 1994 are relatively small (around 9 percent). However, there are a number of problems with these data, namely inconsistencies in the definitions used, the age of collection, and the percentage of the population from whom these data are captured (Ministry of Health, 2002). Moreover, even where disease incidence and potential risk factor values show similar (or opposite) temporal patterns at the national-level, such patterns do not prove causality at the individual-level. As a result, more detailed investigations would be necessary to validate/falsify these possible explanations.

9.2.2 Higher incidence in urban areas

A number of analyses have pointed to an excess of ALL cases in the urban areas of the country. For example, at the CAU-level, the majority of the highest SIRs were located in New Zealand's main cities, often in the city centres. These results were supported by the finding of very low Poisson probabilities in the main urban centres and the identification of significant spatial-temporal clustering of the disease in the urban areas of Christchurch, Tauranga, North Shore and Hamilton. No significant clustering was observed in rural areas. Finally, a number of regression models were supportive of a deficit of ALL cases in rural areas, and ALL was positively associated with population density for the periods 1980-2004, 1980-1991 and 1987-1992. However, a number of exceptions exist: some very high SIRs and low Poisson probabilities were found in rural areas in the CAU-level analyses (although these were usually only based on one observed case). Furthermore there was no significant trend in incidence across the seven urban/rural categories. However, the small number of cases in some of the rural categories may help to explain this finding.

In New Zealand, no previous studies have examined small area urban/rural differences in the incidence of childhood ALL. According to Kinlen's theory, childhood ALL would be expected to be higher in isolated rural areas, following a sudden increase in population mixing. Children living in rural areas are thought to have a lower level of natural immunity to many infections,

including a specific leukaemogenic infection (Kinlen, 2000). However, in New Zealand, ALL was generally higher in urban areas. Moreover, research in other settings shows mixed results. For example, some UK studies have found higher incidence of ALL in rural areas (Alexander et al., 1990, Dickinson and Parker, 1999, Langford and Bentham, 1993), while others have found higher incidence in more densely populated areas (Gilman and Knox, 1998, McNally et al., 2003). Furthermore, studies in Australia (McWhirter and Bacon, 1980), Taiwan (Li et al., 1998), Sweden (Hjalmars and Gustafsson, 1999) and the USA (Adelman et al., 2007, Adelman et al., 2005, Muirhead, 1995) have found a higher risk of ALL in high density urban areas, and thus support the New Zealand findings. While such studies appear to contradict Kinlen's population mixing hypothesis, he has recently suggested that in 'exceptional circumstances' it may be possible to detect excesses of childhood leukaemia following unusual population mixing in urban areas (Kinlen, 2004b, p.716).

It could be argued that in increasingly hygienic and socially segregated societies, children living in urban areas experience reduced exposure to common infections in early life. Such reduced exposure could result from a greater fear of hazards (for example motor vehicle traffic or crime) in urban areas, which could restrict children's outdoor social mixing, especially at young ages. Social or residential segregation in these areas may also play a role in reducing the number of early social contacts of children. These factors, coupled with the generally higher population mixing noted in urban areas in New Zealand (Chapter 6), could mean that when children do eventually come into contact with a variety of infections, they represent a greater challenge to their developing immune systems. Thus, where in the past children living in the most remote and rural areas were likely to have poorly stimulated immune systems, a similar or stronger effect could now be occurring in modern day urban areas in New Zealand.

Alternatively, McNally et al (2003) postulated that a subset of ALL cases could be triggered by infectious exposure *in utero*, and would thus be associated with more densely populated areas, especially at the time of birth. Therefore infections acquired by the mother during pregnancy are presumed to be detrimental in terms of later ALL development. Several case-control studies support this theory (Lehtinen et al., 2003, McKinney et al., 1999a, Naumburg et al., 2002), while others do not (Dockerty et al., 1999d, Infante-Rivard et al., 2000). A second subset of ALL cases may result from delayed exposure to common infections, and would thus be more common in areas of lower population density (McNally et al., 2003). While this theory cannot be tested in this research, it could plausibly explain the inconsistent relationship between ALL and

urban/rural status and the weak positive association with population density. Furthermore, this theory helps to reconcile the contradictory findings of previous studies on this topic.

9.2.3 Higher incidence in the most affluent areas

As well as being higher in urban areas, a raised incidence of childhood ALL was also noted in the most affluent CAUs of New Zealand. However, this result was not compatible with the only previous study on leukaemia and deprivation conducted in New Zealand. A nationwide case-control study was carried out for the period 1990-1993 and showed a significantly increased risk of both combined leukaemias and ALL in children from the 'no paid jobs/not classifiable' (lowest) social class group. While non-participation bias did not appear to affect this finding (Dockerty et al., 1999d), more detail would be required on the jobs which were not 'classifiable' and included in the lowest class category. In general, case-control studies which used interviews or questionnaires to obtain individual-level measures of family income, or mother's or father's education have shown negative associations with childhood leukaemia (e.g. Dockerty et al., 1999d, Ma et al., 2002b, McBride et al., 1999, Poole et al., 2006, Shu et al., 1999), and are thus at odds with the area-level findings of this research. However, studies more comparable to the current research, using ecological designs and measures of average occupational class have observed predominantly positive associations with childhood leukaemia (e.g. Adelman et al., 2007, Alexander et al., 1990, Borugian et al., 2005, Draper et al., 1991, McWhirter, 1982, Poole et al., 2006). Higher incidence of leukaemia was generally observed in the most affluent areas, as found in this research.

In general, lifestyles in affluent societies are thought to insulate infants from common infections before their inevitable exposure through social mixing (Greaves, 2007). As noted previously, a delayed exposure to common infections is thought to be a key element in the aetiology of ALL (Greaves, 2006, Kinlen, 2004a). At the individual-level, better hygiene, fewer other children, reduced breast-feeding or use of crowded child-care facilities, could all contribute to delayed contact with common infections for affluent children. Higher maternal age at child bearing and increased exposure in the home to EMFs from electrical appliances have also been proposed as a link between social deprivation and childhood leukaemia (Little, 1999, UKCCS Investigators, 2000).

Alternatively, it has been suggested that many of the associations found between childhood leukaemia and socioeconomic measures are artefactual (Kuehni and Zwahlen, 2006, Smith et al.,

2006). Issues such as case ascertainment and participation bias plague the interpretation of many of the previous studies. For example under-registration of cases may occur more frequently in poorer neighbourhoods, thus accounting for lower incidence in these areas (Smith et al., 2007). However, case ascertainment within the New Zealand Cancer Registry is thought to be complete (Dockerty et al., 1997), therefore this issue cannot explain the lower incidence observed in the more deprived areas of New Zealand.

A final possibility is that the relationship between deprivation and childhood ALL was confounded by other variables. The most likely candidate is ethnicity. There is a significant disparity in the distribution of deprivation in New Zealand; over half of the Māori population live in the most deprived deciles (Blakely et al., 2002, Robson et al., 2007). Consequently, the finding of a higher ALL risk in the most affluent areas could be due to these areas having a higher proportion of European residents. As previously noted, children of European descent are more susceptible to developing ALL compared to children in the other main ethnic groups in New Zealand (Chapters 3 and 7). Thus, to confirm a real association between ALL and deprivation, future analyses would also need to adjust for the effects of ethnicity. In addition, population mixing could be confounding the relationship between childhood leukaemia and social class. Less deprived areas are often associated with a high number of in-movers, and families of higher social class tend to be more mobile. As a result, studies which have found raised incidence of leukaemia in more affluent areas may have been indirectly observing a population mixing effect (Stiller and Boyle, 1996). However, the relationships noted in this research remained significant after various measures of population mixing were added to the models.

9.2.4 Higher incidence in areas where population mixing increased

Higher incidence of childhood ALL was found in areas which increased the most in population mixing. This central theme was supported by the results of several analyses, for a number of different time periods. The most compelling evidence was from the 6/7 year multivariate regression analyses, where significant positive associations were noted between ALL and a number of different population mixing change measures. Moreover, these associations remained after adjustment for potential confounding variables, including the ethnic composition of the population, area-level deprivation, and population density. However, there was one notable exception to the positive associations found. In the 12/13 year regression analyses, ALL was significantly inversely related to the change in the percentage of child migrants' variable

between 1992 and 2004, although the effect was relatively small and was only significant in two models.

The findings of previous studies on childhood leukaemia and population mixing are mostly compatible with the results of this research. The general picture is of raised childhood leukaemia in areas which have witnessed large increases in population mixing over relatively short time periods. For example, several studies have found a raised risk of childhood leukaemia in areas which have increased considerably in population (Alexander et al., 1997, Clark et al., 2007, Kinlen, 1988, Kinlen et al., 1990, Koushik et al., 2001, Langford, 1991, Wartenberg et al., 2004). In this study, ALL was also significantly higher in areas where the total population had increased the most between 1993 and 1998. Other research has found childhood leukaemia incidence to be associated with increases in commuting (Kinlen et al., 1991), temporary workers (including military personnel) (Boutou et al., 2002, Kinlen, 2006, Kinlen and Balkwill, 2001, Kinlen et al., 1995, Kinlen and Hudson, 1991, Kinlen et al., 1993), and wartime evacuees (Kinlen and John, 1994) and refugees (Labar et al., 2004). While this study did not employ these same measures, it found ALL to be higher in areas where the percentage of total, child or overseas migrants or one year movers had increased the most. In addition, this study was the first to test the potential association between childhood ALL incidence and increases in tourist flows, with significant positive results observed. However, unlike this research, the majority of the previous studies concentrated on examining the effects of population mixing in isolated rural populations. Of the studies which have examined urban areas as well, most have found increased incidence of childhood leukaemia in areas with high or increased population mixing levels (e.g. Stiller and Boyle, 1996), with one study only noting this association in urban areas (Dickinson et al., 2002).

Only four studies contradict the ALL and population mixing findings of this research. Three studies in the UK have found incidence of ALL to be significantly lower in areas of high population mixing at the time of diagnosis (Feltbower et al., 2005, Law et al., 2003, Parslow et al., 2002). The only previous study on population mixing and childhood leukaemia in New Zealand, found no association between population growth and the disease in three small areas of the North Island (Dockerty et al., 1996a). As noted earlier, this study was geographically limited and based on a small number of cases.

As detailed in Chapter 2, differences in study design and population mixing measurements are likely to explain, at least in part, the contradictory results on this topic. In this research, change over time in population mixing was measured over periods approximately similar to the

diagnosis periods. In contrast, Law et al (2003) measured population mixing in 1991 as the proportion of immigrants moving to the census area in the previous year and the diversity of their regions of origin. These snapshot measures of population mixing were then related to ALL registrations for the period 1991-1996. The results showed elevated risks of ALL in areas with a low diversity of origins of migrants. As the authors suggest, this finding is consistent with the Greaves hypothesis, that low infectious exposure in early life could predispose children to development of ALL. A slightly different design was adopted by Parslow et al (2002), who related a snapshot of population mixing in 1991 to childhood leukaemias diagnosed between 1986 and 1996 in Yorkshire, as did Feltbower et al (2005) to cases of ALL diagnosed between 1986 and 1998. Thus in these studies, population mixing was measured at the approximate mid-point of the diagnosis period, making the results not directly comparable with the findings of this research.

The finding of significantly elevated ALL in areas which increased the most in population mixing in this study is compatible with the expectations noted in Chapter 1. It was thought that areas which started with relatively low population mixing would have a pool of children whose immune systems had not been sufficiently challenged. Where increases in population mixing occurred in these areas over relatively short periods of time, there would be an increase in the number and range of infections circulating in the area. Thus, children would be more likely to come into contact with these infections: one of which could be a yet-to-be-identified leukaemogenic virus (following Kinlen's hypothesis) or the stress of any one or more non-specific infections could trigger ALL (following Greaves' hypothesis).

However, it should be noted that the effects of population mixing on ALL incidence in New Zealand (measured using IRRs) were relatively small. So too was the total variation in the distribution of ALL that was accounted for in the models (R^2 values). These findings suggest that additional risk factors are important in the aetiology of childhood ALL. Indeed, population mixing could be working in conjunction with a number of other risk factors (that could not be controlled for here) to create increases in ALL incidence. For example, individual-level factors such as the number of siblings, day-care attendance, family hygiene and social isolation could act to exacerbate a low infectious exposure in early life. Subsequent increases in population mixing could then have a more dramatic impact on a child's immune system, resulting in an abnormal response such as ALL. This theory could help to explain the apparent urban excesses of ALL cases in New Zealand, since individual factors which reduce early exposure to infection could be more prevalent in urban areas. However, this theory would require further investigation.

9.3 Type 1 diabetes key findings

Several prominent themes relating to the geographical epidemiology and potential aetiology of childhood type 1 diabetes in the Canterbury region of New Zealand are also of interest. First, the incidence of type 1 diabetes in children increased substantially between 1980 and 2004 in the area. Second, type 1 diabetes incidence was significantly higher in satellite urban communities and rural areas with a high urban influence. Third, significantly raised incidence of the disease was observed in the more affluent areas of the region. Finally, population mixing was significantly associated with childhood type 1 diabetes in the Canterbury region. Type 1 diabetes was found to be significantly higher in areas with high levels of population mixing measured just prior to the diagnosis period. Moreover, and compatible with the ALL findings, type 1 diabetes incidence was significantly raised in areas where population mixing had increased the most during the study period.

9.3.1 Increase over time

There was a six-fold increase in the incidence of type 1 diabetes in Canterbury children during the 25 year study period. This effect remained after controlling for the age- and sex- specific population at risk, and age group at diagnosis. The results are consistent with a previous analysis of temporal trends in type 1 diabetes incidence in the Canterbury region for the period 1970-1999 (Willis et al., 2002b). The continued rise in type 1 diabetes incidence in Canterbury is consistent with increases observed in Europe, North America, South America, Africa and Asia (DIAMOND, 2006). In Oceania, significant increases in childhood type 1 diabetes have also been observed in Australia (Haynes et al., 2004, Taplin et al., 2005). In Finland, the highest incidence in the world (45 cases per 100,000 population) was recorded in 1996 (Tuomilehto et al., 1999). The highest incidence in the Canterbury region was 32.5 cases per 100,000 population and occurred in 2004.

One explanation for the recent increases in type 1 diabetes that has been suggested is the improvement of case ascertainment over time. However, this research used data from the Canterbury Diabetes Register which is believed to be complete from as early as 1970 (Willis et al., 2005, Willis et al., 2002b). Considering what is already known about the causes of the disease (Chapter 4), possible explanations for the increase could include a growth in the genetically susceptible population, and/or an increase in an environmental trigger for the disease.

It is possible that an increase in the genetically susceptible population could be partly responsible for the rise. Between 1991 and 2001 for example, the total number of Europeans resident in Canterbury increased by over 50,000 people. In Finland, better survival of type 1 diabetes patients in combination with an increase in the number of offspring in patients with type 1 diabetes, especially women, has been proposed to explain increases in the genetically susceptible population (Tuomilehto et al., 1999, Tuomilehto et al., 1995). The same could be true in Canterbury (Willis et al., 2002b). However, in the UK there is also evidence to suggest that the contribution of high-risk susceptibility genotypes has been falling over time (Gillespie et al., 2004), implicating an increasing role for environmental risk factors.

Changes in the prevalence of non-genetic/environmental risk factors could also partly explain the increase in childhood diabetes in Canterbury. Previous work in the Canterbury region (Willis et al., 1997) tested the theory that the introduction of universal hepatitis B vaccination in 1989 caused a 60 percent increase in the incidence of type 1 diabetes (Classen, 1996). However, the study found no evidence to support these claims or to suggest that immunisation at birth had a protective effect (Willis et al., 1997). Therefore other theories require consideration.

A theory yet to be investigated in the New Zealand setting, involves vitamin D deficiency. There is increasing evidence to suggest that vitamin D is protective against type 1 diabetes development (Brekke and Ludvigsson, 2007, Greer et al., 2007, Hypponen et al., 2001, Mathieu and Badenhoop, 2005). It is possible that vitamin D deficiency in New Zealand has increased as a result of campaigning messages by health authorities to encourage seeking shade, covering up and wearing ultraviolet sunscreen lotions to lower the risk of skin cancer (Bulliard and Reeder, 2001, Pearce et al., 2006a). Sunscreens with a sun protection factor of 15 can reduce the capacity of the skin to produce vitamin D by about 98 percent (Shrapnel and Truswell, 2006). In New Zealand, the recent National Children's Nutrition Survey found that 31 percent of children aged 5–14 years were classed as Vitamin D insufficient, and four percent classified as vitamin D deficient (Rockell et al., 2005). No longitudinal data were available.

Weight gain has also been suggested as a causal factor in the pathogenesis of type 1 diabetes (Hypponen et al., 2000, Johansson et al., 1994, Wilkin, 2001). As in many other OECD countries (Couper, 2007), the increases in type 1 diabetes have occurred concomitantly with increases in obesity in New Zealand. Between 1989 and 2003, adult male obesity doubled from ten percent to 20 percent, and adult female obesity increased from 13 percent to 22 percent (Ministry of Health, 2004). 21.3 percent of children aged 5-14 years captured in the 2002 National Children's

Nutrition Survey were overweight and 9.8 percent were obese (Ministry of Health et al., 2003). The only longitudinal data on child obesity trends are from a study in Hawke's Bay which showed a two-fold increase in overweight and a four-fold increase in obese children aged 11−12 years between 1989 and 2000 (Turnbull et al., 2004). No data on childhood obesity were available for the Canterbury region. A study in Finland showed that children who developed type 1 diabetes between 1986-1989 were heavier and taller throughout their childhood compared to children in the control group (Hypponen et al., 2000). This result has been supported by more recent studies in the UK (Betts et al., 2005, Kibirige et al., 2003). However, these studies were unable to test causality between the weight gained and subsequent diabetes diagnosis. Further work in this area is necessary.

As noted for childhood ALL, increases in type 1 diabetes in many countries also parallel a decline in the incidence of common infections (Bach, 2002). The hygiene hypothesis contends that reduced exposure to infections, both postnatally and in early infancy, favours diabetes development in genetically susceptible people (Bach, 2005a, Kolb and Elliot, 1994). Moreover, it relates a decrease in common infections in affluent countries over time, to increases in autoimmune diseases, such as type 1 diabetes (Bach, 2005b). Temporal data on the prevalence of common infections were not available for the Canterbury region. However, decreases such as those seen at the national-level (Population and Environmental Health Group, 2006) are likely.

As with ALL, type 1 diabetes has been inversely associated with breast-feeding (e.g. Borch-Johnsen et al., 1984, EURODIAB, 2002). If changes in breast-feeding rates had a substantial impact on type 1 diabetes, a peak in incidence might be expected to occur at around the same time as reductions in the breast-feeding rate. However, between 1980 and 1986 there was a general increase in both the percentage of babies being breastfed and type 1 diabetes incidence. A trough in type 1 diabetes incidence between 1987 and 1988 is consistent with a small peak in national breast-feeding rates in 1987, and a peak in type 1 diabetes incidence is consistent with a small trough in breast-feeding rates in 1990/1991. Furthermore, type 1 diabetes in Canterbury has increased steadily since 1995 and national breast-feeding rates have been decreasing steadily since 1994 (Ministry of Health, 2002). However, the peaks and troughs noted do not account for a lag time between lows in breast-feeding occurrence and a subsequent peak in type 1 diabetes incidence; depending on the age at diagnosis, there could be a lag time of up to 14 years. More importantly, these data are based on average rates for the whole of New Zealand, and different breast-feeding patterns may be evident in Canterbury, and especially at the individual-level.

9.3.2 Higher incidence in satellite urban communities and rural areas with a high urban influence

Results from the regression analyses revealed significantly higher incidence of type 1 diabetes in satellite urban communities compared to main urban areas, for the periods 1980-2004, 1992-2004 and 1999-2004. These areas are towns and settlements with strong links to urban centres and where 20 percent or more of the population work in a main urban area (Statistics New Zealand, 2006a), in this case Christchurch. In addition, significantly higher incidence occurred in rural areas with a high urban influence compared to the Christchurch main urban area in 1980-2004 and 1999-2004. These CAUs form a transition area between the Christchurch main urban area and the rural areas beyond. A significant proportion of the population living in these areas, work in Christchurch (Statistics New Zealand, 2006a). Further evidence of higher type 1 diabetes incidence in these areas can be seen in the results of the cluster analyses. Two of the significant clusters observed included the satellite urban communities of Rangiora, Southbrook and Woodend, and the rural areas with a high urban influence of Waikuku and Tuahiwi. The two clusters covered the same area and time period (1998-1999) and occurred in male children aged 0-14 years and 5-9 years at diagnosis.

Higher incidence of the disease was thus noted in areas just outside and within commuting distances of Christchurch, especially in the latter half of the study period (1992-2004). This finding is consistent with increased levels of counter-urbanisation in the region; more people are choosing to live in less densely populated areas away from the City of Christchurch (Buchanan, 2004). The majority of these people are unlikely to be taking up farming lifestyles and thus their children miss out on the protective effects of frequent contacts with cattle, poultry and cats (Johnston and Openshaw, 2001), and also more frequent human contacts associated with main urban areas (Anderson and May, 1982, Cox, 2007). Alternatively, this result could suggest that the increases in population noted in these areas have brought in new infections to which some children have had an abnormal immune response. For example, during the period 1996-2001, the Rangiora/Waikuku/Camside cluster area witnessed a 35.5 percent increase in child migrants. It is plausible that an influx of new children to the area brought with them infections which acted as a trigger to diabetes onset, and resulted in an excess of cases diagnosed between 1998 and 1999. If this theory is correct, one would expect a decrease in type 1 diabetes incidence in these areas in the future if the population continues to rise and reaches some critical threshold level.

The high or increased levels of commuting in these areas could also be responsible for the higher than expected incidence of type 1 diabetes. This study did not investigate the effects of commuting on type 1 diabetes incidence as previous studies for childhood leukaemia showed mixed results (Kinlen et al., 1991, Stiller and Boyle, 1996), and parental occupational exposure was not significantly associated with type 1 diabetes in their offspring in a case-control study in Yorkshire and Northern Ireland (Fear et al., 1999). However, the urban/rural findings from this study suggest that the importance of differential commuting levels should be explored. Interestingly, the only independent urban community in the study region (Hanmer Springs) where less than 20 percent of the population work in a main urban area, had significantly lower incidence of type 1 diabetes. Since Hanmer Springs is 135km (1.5 hours drive) away from Christchurch, few residents are likely to commute daily between the two areas.

9.3.3 Higher incidence in more affluent areas

Higher incidence of childhood type 1 diabetes was also observed in the most affluent areas of the Canterbury region. Supporting evidence was found in the analysis of both standardised incidence ratios by deprivation deciles and quintiles, and incidence rate ratios derived from Poisson and zero-inflated Poisson regression models. However the relationship was not consistently significant and could suggest that the relationship was weak, or that the association was only important in certain time frames. Alternatively, since the deprivation score was created using 2001 census data (Salmond and Crampton, 2002a, Salmond and Crampton, 2002b) it may not be a reliable representation of deprivation levels for earlier time periods.

The Canterbury findings are consistent with the majority of previous studies. For example, an inverse relationship between deprivation and type 1 diabetes incidence has been observed in Western Australia (Haynes et al., 2006), England (Feltbower et al., 2005, Parslow et al., 2001), Scotland (Patterson et al., 1994, Patterson and Waugh, 1992), Northern Ireland (Cardwell et al., 2006, Patterson et al., 1996, Patterson et al., 1994) and at the national-level in Europe (Patterson et al., 2001, Tedeschi and Airaghi, 2006). However, an older study of Northern England (Crow et al., 1991) and a recent study in Germany (du Prel et al., 2007) found a positive relationship between the disease and deprivation. Other studies (Baumer et al., 1998, Cox, 2007, Evans et al., 2000) have found no association. No previous studies have examined this issue in New Zealand.

There are a number of possible explanations for the higher incidence of type 1 diabetes observed in affluent areas in this study. As considered with childhood ALL, ethnicity may be confounding

the relationship between type 1 diabetes and deprivation, and needs to be controlled for in future analyses. Where the association remains after adjustment for ethnicity, area measures of social deprivation may be acting as a proxy for individual and/or area-level risk factors for childhood type 1 diabetes. At the individual/household-level a number of potential risk factors which vary by socioeconomic status (SES) have been linked to the aetiology of type 1 diabetes. Such risk factors include; diet, breast-feeding practices, maternal age, family size, household overcrowding, hygiene levels and childcare practices (du Prel et al., 2007, Haynes et al., 2006). For example, increased household hygiene and smaller family sizes, often associated with more affluent families, are likely to result in lower infectious exposure in childhood and thus could be responsible for the high incidence of type 1 diabetes in these families (Bach, 2002, Kolb and Elliot, 1994). However, a limitation with using an area-level deprivation measure is that the SES of individual children who developed type 1 diabetes remains unknown. Less affluent children living in affluent areas could represent the majority of children suffering from type 1 diabetes in these areas.

Risk factors could also operate at the area-level. For example, more affluent neighbourhoods might be more socially segregated. 41.5 percent of CAUs in the Christchurch main urban area have been found to be affluent and segregated compared to only 27.9 percent which were deprived and segregated (Barnett, 2000). A greater fear of crime in these areas could result in individual houses being surrounded by large walls and fences, and common play areas being restricted to dissuade the presence of people from other areas. These factors may confine the free movement and mixing of children in and around the home. In addition, due to the type of people living in these areas, there may be smaller family sizes resulting in fewer children in general. Lower child population densities have implications for the spread of infections between children. Furthermore such neighbourhoods are usually cleaner and have less rubbish in the streets. All of these factors may act to limit childhood exposure to infections in early life. However, further research would be required to test these ideas.

9.3.4 Higher incidence in areas where population mixing was high or had increased

Since infections have been implicated in both protecting against and triggering type 1 diabetes (Bach, 2005a, Filippi and von Herrath, 2005), two main types of population mixing measures were examined. Static levels of population mixing were assessed at the beginning of each diagnosis period to test whether high levels of infectious exposure in early life protected against subsequent type 1 diabetes development. In addition, changes in population mixing levels were

assessed over time to examine whether large increases in population mixing could act as a trigger for this disease. Two main findings were identified: type 1 diabetes was higher in areas where population mixing levels were high prior to the diagnosis period, and also in areas which had increased the most in population mixing during the diagnosis period. These two themes are now discussed in turn.

Where population mixing levels were high

Population mixing levels were measured at the beginning/prior to the diagnosis period as a proxy for early life exposure to common circulating infections. Since early infectious contact is thought to be beneficial for the immune system and to reduce the chance of autoimmune responses, areas with high population mixing were expected to have lower incidence of type 1 diabetes in the following years. However, the results showed that the opposite occurred. Type 1 diabetes incidence was significantly higher in areas with *high* levels of population mixing measured at the beginning/prior to the diagnosis period. All of the significant findings were in a consistent direction, and this trend was observed regardless of the length of the analysis period. However, this effect was only noted in a sample of the static population mixing measures tested. The percentage of overseas visitors was the only variable found to be a significant predictor of the disease for the 12/13 year time periods. In the 6/7 year analyses, the percentage of overseas visitors and overseas migrants' variables were only significant in one model each. Furthermore, no significant associations between type 1 diabetes and these static population mixing measures were noted after 1992.

The results of three previous population mixing studies do not support these Canterbury findings. Small area analyses in Yorkshire in England (Feltbower et al., 2005, Parslow et al., 2001), and in Tayside in Scotland (Cox, 2007), all found significantly higher type 1 diabetes incidence in areas of low population mixing. The majority of ecological studies which used measures of household overcrowding, deprivation, population density and urban/rural status as proxy measures for infectious exposure, concur with the population mixing findings in the UK (e.g. Cardwell et al., 2007, Feltbower et al., 2005, Haynes et al., 2006, Parslow et al., 2001, Patterson et al., 1996, Staines et al., 1997). Furthermore, case-control studies on infections in early life, birth order, and day-care attendance tend to support the notion that early exposure to common infections is protective against later development of type 1 diabetes (e.g. Haynes et al., 2007a, McKinney et al., 2000, Pundziute-Lycka et al., 2000).

Differences between the designs of the Canterbury study and previous studies are likely to partly explain the varying results. In the Canterbury study, static population mixing was measured at the beginning or just prior to the diagnosis periods, as it was assumed that children would need to be exposed to high population mixing levels for a number of years in early life before it could confer any protection. Furthermore, where this early exposure to common infections is absent and type 1 diabetes is eventually initiated, it is usually a number of years before any symptoms occur and a diagnosis is made. Therefore, the population mixing levels that are likely to be important are those that occurred several years before diagnosis. Both of the studies in the Yorkshire region of the UK measured population mixing at the approximate mid-point of the study period. For example, Parslow et al (2001) modelled type 1 diabetes cases diagnosed between 1986 and 1994 in relation to population mixing levels in 1991. Therefore presumably by 1991 the majority of cases had already been diagnosed, and cannot have been affected by the population mixing levels in 1991. In the study of Tayside in Scotland, type 1 diabetes cases for the period 1998 to 2001 were associated with population mixing in 2001, the final year of the diagnosis period (Cox, 2007). Thus, these studies fail to incorporate sensible lag times between the population mixing measurements and case diagnoses. The population mixing measurements made are thus unlikely to represent population mixing levels in early life for the majority of children in these studies.

Differences in the length of the diagnoses periods between this study and the previous work may also have had a bearing on the contradictory results achieved. The Tayside study analysed data for the shortest length of time (4 years), followed by this study (6/7 years), the Parslow et al study (9 years), this study (12/13 years) and Feltbower et al (13 years). The optimal temporal scale for measuring the effects of population mixing on type 1 diabetes incidence is not yet known. This study is the first to consider the importance of different temporal durations on the results. In this research, consistent results were found over both the 12/13 and 6/7 year analyses.

More specifically, in this research, a higher incidence of type 1 diabetes was noted after high levels of overseas migrants and overseas visitors were recorded. Furthermore, areas with higher than average in-migration and higher than average migrant diversity were also positively related to the incidence of this disease. Therefore the results highlight the importance of *where* new entrants to an area came from. Perhaps due to New Zealand's isolated location, common infections brought in from overseas were detrimental to local children if encountered in early life. However, these variables were only significantly associated with type 1 diabetes incidence in the first half of the study period. It is possible that infections from overseas became less

important as the number of overseas visitors and migrants increased over time, and these infections became more common place in the region. The number of overseas visitors in Canterbury has increased substantially from around 3,000 people in 1981 to over 10,000 by 2001. Thus perhaps after 1992, the majority of children were exposed to these infections (and more frequently) and therefore they represented less of a threat to their developing immune systems. This theory would suggest that some kind of threshold effect exists. However, if infections from overseas ceased to be an important cause of type 1 diabetes cases after 1992, a decrease in the incidence of the disease would be expected after this time. As reported previously, type 1 diabetes incidence increased steadily throughout the 1990s and early 2000s in this region.

Where population mixing had increased

Higher incidence of type 1 diabetes was also observed in areas where population mixing had increased the most over short time periods (6/7 years). However, this outcome was only observed after 1986, and only in some of the continuous population mixing measures. The most convincing results were shown in the regression models for the diagnosis period 1999-2004. During this time, three different population mixing change measures (1996-2001) were all positively associated with type 1 diabetes incidence. Areas which had witnessed an increase in the volume of in-migration (all ages or child) or residential mobility had higher childhood diabetes incidence between 1999 and 2004. Further evidence to suggest an infectious trigger for type 1 diabetes was found in the cluster analyses, with a number of significant clusters noted at the time and place of disease diagnosis. These findings support the hypotheses that higher incidence of type 1 diabetes would be found in areas which increased the most in population mixing. Areas which have low levels of population mixing are thought to have a low prevalence and range of common infections. Therefore children in these areas potentially lack early exposure to infections which are necessary for immune system modulation. Where these areas experience an increase in population mixing, especially over short time periods, new infections may act as the final trigger to diabetes in genetically susceptible children.

To date, no other studies have examined the potential triggering role of increases in population mixing in the aetiology of type 1 diabetes. However, a large number of studies have found excess cases of childhood ALL in areas which have experienced population growth (e.g. Alexander et al., 1997, Kinlen, 1988, Kinlen, 2006, Koushik et al., 2001, Wartenberg et al., 2004), and increases in migration and migrant diversity (e.g. Boutou et al., 2002, Dickinson and

Parker, 2002, Labar et al., 2004, Rudant et al., 2006, Stiller and Boyle, 1996), and are compatible with these diabetes findings. As noted earlier, these two diseases have many epidemiological and potential aetiological features in common. Furthermore, a number of specific viruses (e.g. rubella, mumps, rotavirus and enteroviruses) have been implicated in causing type 1 diabetes (Filippi and von Herrath, 2005, van der Werf et al., 2007) and thus support the conclusions of the Canterbury analyses. Serological evidence and case reports of viral infection preceding type 1 diabetes onset have been noted (Yoon and Jun, 2003). Additionally, epidemiological studies which have found seasonal variation (Lévy-Marchal et al., 1995, Willis et al., 2002a) and significant space-time clustering of type 1 diabetes, are all suggestive of a viral aetiology (e.g. Bodington et al., 1995, Dahlquist and Kallen, 1996, Feltbower et al., 2006, McNally et al., 2006a). Animal models have also identified a number of diabetogenic viruses in mice and rats (Jun and Yoon, 2001). Moreover, the Canterbury findings are not mutually exclusive with the hygiene hypothesis; that low exposure to microbial infections in early life increases the risk of type 1 diabetes (Bach, 2005b, Kolb and Elliot, 1994). These analyses test an extended version of this hypothesis; that where population mixing levels change from low to high over short periods of time, the increased exposure to infections could be the final trigger for type 1 diabetes onset.

Despite support for a positive relationship between population mixing change and type 1 diabetes incidence in the Canterbury region, there were some exceptions. Negative associations were observed between the disease and population mixing change for the diagnosis period 1980-1986. Moreover, in the 12/13 and 25 year analyses, increases in population mixing were associated with a decreased risk of type 1 diabetes for the periods 1992-2004 and 1981-2004.

Inconsistencies between the 6/7 and the 12/13/25 year time periods could arise from different responses to population mixing changes over varying lengths of time. For example, increases in population mixing over longer time periods (13 and 25 years) were found to be protective against type 1 diabetes. This result could suggest that slower increases in population mixing help to build up children's immune systems through more gradual exposures to common infections. Additionally, a shorter and sharper growth in population mixing (6/7 years) may be much harsher on a developing immune system, and more likely to result in an abnormal immune response. It could also be argued that the 12/13 and 25 year time periods are too long to accurately measure changes in population mixing. Considerable fluctuations in population movements and mobility are likely to occur within these time spans.

Irregularities between the first 7 years of the study (1980-1986) and the next three 6 year periods could be explained by the small number of cases noted in the first 7 years. Only 55 new cases were observed in the seven year period between 1980 and 1986, compared to 120 new cases observed during the six years between 1999 and 2004. Perhaps as a consequence of these smaller numbers, only model 5 revealed significantly lower type 1 diabetes incidence in areas with the largest population growth between 1980 and 1986. In addition, the confidence intervals for the change in the percentage of child migrants IRRs included one in every model for this period. Another potential explanation is that increases in population mixing between 1980 and 1986 might not have been substantial enough to result in excess type 1 diabetes cases prior to 1986. Chapter 6 showed that large decreases in the percentage of total and child migrants occurred across New Zealand between 1981 and 1986. This was primarily due to a high number of internal movers in 1981. Similar patterns were found in the Canterbury data. In addition, the final six years of the study allow for a longer lag period between the population mixing changes (1996-2001) and the diagnosis period (1999-2004). In areas where population mixing increased between 1996 and 2001, the high levels of population mixing in 2001 would be expected to be the potential trigger for beta cell destruction. Depending on the child, it could take months to years before any symptoms appear and a diabetes diagnosis is made. Therefore an increase in type 1 diabetes could be observed a number of years after the increase in the potential environmental trigger (2001). The diagnosis period of 1999-2004 takes a lag time of up to 4 years into consideration.

Thus in general, the incidence of type 1 diabetes was higher in areas with high population mixing prior to the diagnosis period, or in areas which increased the most in population mixing throughout the diagnosis period. The principal findings of the type 1 diabetes analyses therefore support a triggering, rather than protective effect, for infections introduced through population mixing, especially over short periods of time.

9.4 Comparison between ALL and type 1 diabetes

There were a number of noteworthy similarities between the ALL and type 1 diabetes findings. First, the rates of both diseases were higher in children of European ethnicity. Even after controlling for the population at risk, incidence was much lower in children from other ethnic groups. This occurrence is probably due to the higher genetic susceptibility noted in European populations for both diseases and is consistent with studies in other settings (Karvonen et al., 1997a, Liang and Pui, 2005).

Second, both diseases increased significantly during the 25 year study period (1980-2004). However, type 1 diabetes increased more substantially than ALL. Differences between the rates of increase in the diseases could be explained by the lack of comparability between the national ALL rates, and regional type 1 diabetes rates. The national incidence of ALL is likely to hide much regional variation, and type 1 diabetes incidence in Canterbury has previously been shown to be high in comparison with other regions in New Zealand (Campbell-Stokes and Taylor, 2005, Willis et al., 2004). It is also plausible that different causal mechanisms are driving the rise in the diseases. However, a number of common explanations for these increases are conceivable, including a reduction over time in protective factors such as the prevalence of common childhood infections or breast-feeding. Moreover, an increase in social segregation and/or a decline in children's use of public space could have contributed to a general decrease in children's actual exposure to common infections. Such hypotheses are consistent with the hygiene hypothesis.

Third, and also compatible with the hygiene hypothesis, was the observation that both diseases were inversely associated with area-level deprivation. Affluence could be a proxy for individual and/or area-level risk factors that have been implicated in the aetiology of both diseases. For instance, at the individual-level such risk factors include breast-feeding, birth order, childcare, and increased hygiene (Bach, 2001, Greaves, 2006, Kolb and Elliot, 1994). Several area-level risk factors can also be postulated, including increased social segregation, restricted play of children, lower child population densities and cleaner neighbourhoods.

Fourth, the incidence of both ALL and type 1 diabetes was generally lower in the rural CAUs of New Zealand/Canterbury. For example, the IRRs for the three most rural categories were below one in the majority of models for both diseases. Furthermore, significantly lower incidence was noted for both diseases in areas classed as rural with a low urban influence (1980-2004). The number of cases in the most rural areas (classed as highly rural/remote) was very low for both diseases (ALL = 12, type 1 diabetes = 8) and probably accounts for the lack of statistical significance in the models. At the urban end of the spectrum, the highest IRRs were noted for satellite urban communities and rural areas with a high urban influence for both diseases. However, these IRRs were only significant in the type 1 diabetes analyses. Interestingly, the directions of association for population density were different: ALL was positively associated with population density, and type 1 diabetes was inversely associated with population density.

Finally, both ALL and type 1 diabetes were significantly higher in areas which had increased in population mixing over short time periods (6/7 years). Moreover, similar population mixing measures were implicated in explaining the variation in the two diseases. For example, both were found to be higher in areas which had increased the most in: total population, the percentage of total migrants, the percentage of child migrants, the percentage of overseas visitors and the one year mobility percentage. However, these significant relationships did not necessarily occur during the same time periods, and of course, relate to population mixing in either the whole of New Zealand or just the Canterbury region. Interestingly though, the incidence of both diseases was significantly raised in 1999-2004 in areas which had increased the most in the one year mobility percentage. In addition, neither disease was significantly related to the change in the migrant diversity variable.

9.5 Critical assessment

There are a number of potential limitations to this study which should be acknowledged relating to the data and methods of analysis used.

9.5.1 Data issues

Variations in the accuracy and detail of the address at diagnosis data collected for each disease may have affected the findings of this research. This problem was considerably worse for the ALL data obtained from the New Zealand Cancer Registry (NZCR). Firstly, due to missing addresses in the NZCR for the period 1993-2001, address data from two other sources were also used (Chapter 5). Unfortunately, addresses from these sources were not necessarily recorded at the time of the ALL registration, and thus may not represent the address where every child developed ALL. For 77 percent of the total ALL cases, the address used for geocoding had been recorded within two years of the ALL registration date, and was thus likely to represent the true address at diagnosis. However, the exact number of cases which this problem affects is unknown, and therefore the implications for the findings of this research are unascertainable.

Secondly, regardless of which data source was used, some addresses for both diseases lacked sufficient information for an accurate house, street or suburb match to be made. Approximately five percent of the total ALL cases and five percent of the total type 1 diabetes cases were only accurately matched to the town/city-level and in this instance were placed in a CAU in the appropriate town/city centre. This misplacement could have affected any of the geographical

analyses at the CAU-level or smaller. For example, some cases which were erroneously placed in central city areas may have belonged in suburb areas near to other cases. As a result, some significant spatial-temporal clustering may have been missed. When conducting spatial analyses on rare diseases, moving one or two cases into different CAUs could have important affects on the findings. Moreover, the significant clustering identified by the analyses might depend upon erroneously placed cases. Due to the small number of cases involved in all of the clusters found, the accuracy of the geocoding for each case was checked. All of the type 1 diabetes clusters were based on cases which were matched to the correct house address. In the ALL analyses, cases involved in the Christchurch, Tauranga and North Shore clusters were all based on exact matches. However, the significant cluster of two 0-4 year olds identified in a meshblock in central Hamilton during 1983 was based on two town/city matches and can thus be considered erroneous. While this placement inaccuracy influenced where exactly a case was placed within a town/city, it should not have affected its urban/rural classification. The suburbs of all towns and cities in New Zealand have the same urban/rural classification as the town/city centre areas. As a result, geocoding errors alone are unlikely to account for the urban excess of ALL cases in New Zealand.

Thirdly, rural addresses were more likely to be inaccurately placed compared to urban addresses. Only 33 percent of the rural ALL cases were matched to an exact house address, compared to 92 percent of urban matches. For the type 1 diabetes cases in Canterbury, these figures were 37 and 93 percent respectively. Studies in the USA have also found that geocoding is generally more successful in urban compared to rural areas (e.g. Kravets and Hadden, 2007). This urban/rural bias in accuracy may have affected the urban/rural SIR and regression analyses, especially due to the small number of total cases resident in rural areas. However, CAUs of similar urban/rural status tend to be located next to each other, thus minimising this bias. Moreover, this problem only affected five percent of the total cases for each disease and was thus unlikely to have had a large impact on the overall findings of this study.

9.5.2 Methodological issues

General issues

A number of general methodological issues could also have affected the results of this study. First, the potential effects of the modifiable areal unit problem (MAUP), whereby different results are obtained from analysis of the same data grouped into different areal units, should be

considered. Different results could be achieved depending on the location of the boundaries imposed on the data, and also the differing scales of analysis employed (Flowerdew et al., 2008, Fotheringham and Wong, 1991, Manley et al., 2006, Openshaw, 1984). Consequently, it is valuable to carry out the analyses for different zones and scales to test the robustness of the results (Haynes et al., 2007b, Manley et al., 2006, Rezaeian et al., 2006). Unfortunately the effects of the MAUP could not be tested in this study since these population mixing data were only available for the area units provided by Statistics New Zealand, and were not available for smaller geographical areas (Meshblock-level). These data could have been aggregated to larger (Territorial Authority) areas but such areas are too extensive to provide meaningful proxies of childhood infectious exposure. The chosen unit of analysis, census area units (CAUs), are aggregations of meshblocks which generally follow features that are identifiable on the ground (road patterns, river courses etc) (Statistics New Zealand, 1997), so both geographies have been designed with some consideration of sociological features.

Second, this study relies upon the untestable assumption that individual-level exposure correlates with area-level measurements (termed the ecological bias or fallacy) (Rothman and Greenland, 1998). Variation in population mixing levels within areas would mean that not all children were exposed to the average population mixing level/changes assigned to an area. However, the population mixing variables used in this study have no appropriate individual-level equivalent.

Third, residual confounding may have occurred. A number of individual-level variables previously associated with these diseases (for example breast-feeding, diet, birth-weight, and chemical and radiation exposure) could not be controlled for in this study. At the area-level, social segregation, which could have affected children's levels of exposure to population mixing, was also not adjusted for. Unmeasured confounding could have under- or over-estimated, or even reversed the direction (McNamee, 2003, Smith and Phillips, 1992) of the associations between population mixing and the diseases.

Fourth, population mixing was only a proxy measure for the number and range of infections entering an area. The variables used were unlikely to pick up on any short-lived epidemics of infection that occurred, and which may be important in explaining increases in the diseases over small spatial and temporal scales, especially significant clusters. Moreover, collinearity between the population mixing measures is very likely since the population change, total migrants, child migrants, overseas migrants, and one year mobility percentage variables will contain some of the same people.

Finally, this research was also unable to account for changes in the genetic structure of the population at risk. Both diseases are known to be higher in children of European descent. However, the population mixing measures do not distinguish between the ethnicity of the new people entering each area. Thus if the majority of new migrants are non-European (and therefore less susceptible to ALL and type 1 diabetes) this would dilute the pool of high risk European populations.

Population mixing measures

The population mixing variables which measure the percentage or percentage change of migrants (total migrants, child migrants, overseas migrants), relate to residential relocation that occurred during the five years preceding each census year. As a result, the movements of people who relocated multiple times within the five year period would be missed. Fortunately, the New Zealand gathers additional information on the length of time individuals have spent at their current residence. These data were used to create one year mobility percentages and thus to capture some shorter term movements. However, this measure will also include moves that occurred within the same CAU which may be less important in terms of the introduction of new infections. The percentage of overseas visitors present on census night was employed to capture some shorter term, but longer distance movements. However, it should be noted that this variable is just a snapshot of tourists as of March every census year, and their numbers and geographical distributions may vary throughout the rest of the year.

Other more temporary movements were not captured by this research. For example, previous studies on ALL have investigated the associations between commuting patterns and disease rates. A significant excess of leukaemia cases was found in towns in the decile with the greatest increase in commuting in one study of West Berkshire in the UK (Kinlen et al., 1991). However, no associations were noted in a more comprehensive recent study of districts in England and Wales (Stiller and Boyle, 1996). Commuter movements have generally been thought to be more important for adult population mixing (Parslow et al., 2001) and were thus not considered here. However, the results of the type 1 diabetes urban/rural analyses in Canterbury showed higher incidence in areas with high levels of commuting, so this aspect of population mobility is worth exploring. In addition, other regular movements more relevant to children, such as travelling to preschool and school, going with parents and caregivers to supermarkets, or to visit friends and family, were not taken into account in this research. It was felt that such movements were likely

to be over shorter distances and their more regular nature would be unlikely to represent a considerable threat to developing immune systems.

A more general issue encountered relates to determining change in the population mixing measures over time. This study measured change as the difference between the start and end year, relative to the start year, and was considered an adequate way to identify areas which experienced increases in population mixing over time. However, it is still unclear how much of an increase is necessary to represent an immunological threat and trigger childhood ALL or type 1 diabetes. Is there a threshold effect? Furthermore, does this threshold level vary between different types of area (e.g. urban/rural settings), or areas which contain different types of people? For example in some rural CAUs, the increase in the raw number of movers was quite small, but since these areas had small populations to begin with, these increases may have been enough.

If population mixing levels increase from high to higher, no excess of ALL or type 1 diabetes cases was expected since children in these areas should have relatively developed immune systems. This study used population mixing categories in order to identify these areas. However, using the national average as a cut off point to determine whether areas have high or low migrant diversity and high or low in-migration is arbitrary. For example, in 2001 the mean percentage of total migrants in New Zealand CAUs was 50.43 percent. Thus, in CAUs classed as low migration areas migrants could still account for up to 50 percent of the total population. This problem also affects the population mixing change categories, which were created from the original classifications of low/high migrant diversity and in-migration.

When designing this study, estimates of the length of time between population mixing exposures and disease diagnoses were made. These estimates tried to take into account both the length of time that might be needed for a population mixing exposure to affect a developing immune system, and the latency period between disease initiation and diagnosis. Little discussion regarding the first part of this equation has been made in the literature to date. In terms of the second part, variable latency periods have been suggested for both ALL (Wartenberg et al., 2004) and type 1 diabetes (Petrovsky and Schatz, 2003). As a result the timing of the population mixing measurements used in this study may not be capturing the protective/triggering aspects of this phenomenon as well as hoped.

The problem of appropriate lag times is compounded by the effects of migration. The longer the time lag between the population mixing measurement and each individual diagnosis, the more likely migration is to have occurred between the two. This problem is especially likely to affect the 12/13 year analyses. The results for the 6/7 year analyses should be more reliable. Moreover, the in/out migration would have to be biased in one direction to have a considerable effect on the results. The issue of migration between exposure and diagnosis is a problem inherent in all cross-sectional studies. In addition, the use of different population mixing and diagnosis periods could have resulted in different results. Furthermore, high levels of population mixing/infectious exposure are most important in early life. As a result, the child's age at diagnosis is also important, but was not properly investigated in this research.

9.6 Conclusion

Despite the limitations of the study, this chapter has discussed a number of important findings regarding the geographical epidemiology and potential aetiology of childhood ALL and type 1 diabetes in New Zealand. Both ALL and type 1 diabetes were found to be more common in children of European descent compared to children belonging to Māori, Pacific and Asian ethnic groups. These findings are consistent with previous New Zealand studies, and the international literature, and are likely to represent an increased genetic susceptibility in European children. Second, both diseases increased significantly during the 25 year study period, and parallel increases noted in other developed countries. A number of potential explanations were put forward to explain these increases, some of which were compatible with an infectious cause and which were plausible for both diseases. Third, both childhood ALL and type 1 diabetes incidence were significantly higher in the most affluent neighbourhoods of New Zealand and Canterbury respectively. This trend lends support to the hygiene hypothesis since lifestyles in affluent areas are thought to insulate children from early exposure to infection. Fourth, the incidence of both diseases was found to be lower in rural areas of New Zealand/Canterbury. This finding confirms the assertion made in this thesis that population mixing should not just be considered in remote rural areas of the country. Finally, a distinct geography of incidence and a number of significant space-time clusters were observed for both diseases, and are suggestive of an infectious trigger for both diseases.

Also indicative of a role for infections in the aetiology of ALL and type 1 diabetes, are the findings that an increase in population mixing over short time periods (6/7 years) was generally associated with an increased risk of both ALL and type 1 diabetes. These findings were as

expected, and compatible with the majority of previous findings. In addition, there was some evidence to suggest that high population mixing levels in early life increased the risk of later type 1 diabetes development. The few previous studies on this topic do not support this result. However, this finding was only noted prior to 1992 and was only supported by a few models. Thus overall, high or increased levels of population mixing were found to be associated with increased risk of ALL or type 1 diabetes within New Zealand. The final chapter of the thesis will draw together the prominent themes of this research and will discuss how these themes can be developed further in future studies.

Chapter 10 - Conclusion

10.1 Introduction

This final chapter summarises the key themes and contributions of this thesis and presents a number of future avenues of research. The chapter begins by briefly revisiting the aims of the study, and how these aims were addressed. The subsequent section discusses two central themes identified by this research. Finally, dissemination of the results is considered, followed by a discussion of a number of future research directions.

10.2 Study purpose

The purpose of this study was two-fold: to examine the geographical epidemiology of childhood ALL and type 1 diabetes within New Zealand; and to test whether the incidence of either disease was associated with small area population mixing. These aims were considered important for a number of reasons. First, both diseases are increasing in many affluent countries, but their exact causes remain unclear. Second, population mixing has been implicated in the aetiology of both diseases, but previous studies show mixed results and population mixing itself tends to be under-theorised. Finally, this issue has not been adequately assessed in New Zealand, a country characterised by increasing population mobility.

To address both aims, a range of geographical approaches were employed. For example, the geographical epidemiology of each disease was determined by the calculation and mapping of age-standardised incidence rates and Poisson probabilities for CAUs, and spatial-temporal cluster analyses were also conducted. Such methods are useful for generating hypothesis regarding possible environmental causes of disease (Lawson, 2001). For the population mixing analyses, Poisson and negative binomial regression models were utilised to assess the associations between small area population mixing and each disease. The findings of these analyses have highlighted two key themes which will each be addressed in turn.

10.3 Key themes identified by the research

First, despite the analyses of type 1 diabetes being confined to the Canterbury region in the South Island of the country, striking similarities were noted with ALL. The majority of the findings were suggestive of an infectious aetiology for both diseases. Second, high or increased levels of

population mixing were generally associated with an increased risk of childhood ALL and type 1 diabetes. Higher incidence of both diseases was observed in areas which increased the most in population mixing over short periods of time (6/7 years). Furthermore, raised type 1 diabetes incidence was also associated with high population mixing in early life.

10.3.1 Similarities between childhood ALL and type 1 diabetes

In addressing the first aim of this thesis, a number of similarities have been noted between the incidence of childhood ALL and type 1 diabetes. Both diseases were found to be higher in the most affluent neighbourhoods, to be lower in rural areas and in areas with a high percentage of household overcrowding, and both had distinct small area geographies and a tendency to cluster at the CAU-level. Moreover, both diseases increased significantly during the study period, and were most common in children of European descent. These findings are compatible with the majority of previous studies on this topic. In addition, some of these findings are supportive of an infectious aetiology for both childhood ALL and type 1 diabetes. One of the several theories put forward to explain the temporal increases in both diseases, relates to a decrease in the prevalence of common infections that has occurred in many affluent countries over recent decades. Decreased exposure to infection in early life is thought to play a crucial role in the development of both diseases. Furthermore, at the individual-level, affluence has been associated with lifestyles which insulate children from early infectious exposure. At the area-level, increased social segregation reduced outdoor play and lower child population densities could limit early childhood infectious contacts in affluent areas. However, further work would be required to empirically test these hypotheses. Furthermore, the potential confounding effects of ethnicity need to be assessed. The finding of significant spatial-temporal clusters of each disease can also be considered as indirect evidence implicating infections in their aetiologies. In addition, the significant associations found between each disease and area-level population mixing, a proxy for infectious exposure, is also indicative of an infectious cause for childhood ALL and type 1 diabetes.

10.3.2 High/increased population mixing levels as detrimental for child health

Increases in population mixing over short periods of time

This study adds to a considerable body of evidence implicating increases in population mixing in the aetiology of childhood leukaemia. Moreover, it contributes to a small but growing area of

research on the effects of population mixing on childhood type 1 diabetes. Increases in population mixing, especially over short time periods (6-7 years) were found to be associated with an increased risk of both childhood ALL and type 1 diabetes. This finding was noted for each 6/7 year time period in the ALL analyses (1980-1986, 1987-1992, 1993-1998 and 1999-2004), and for all but the earliest 7 year time period (1980-1986), in the type 1 diabetes analyses. These relationships remained significant after controlling for the ethnic composition of the population, population density, and deprivation; all factors which have been associated with these diseases at a small geographical scale in other settings (e.g. Parslow et al., 2001, Staines, 1996). Moreover, the findings were not limited to remote rural areas as tested by some earlier studies (e.g. Kinlen et al., 1993, Kinlen and Petridou, 1995). No previous research has examined the importance of increases in population mixing for childhood type 1 diabetes, whereas the ALL results of this research are consistent with the majority of previous ALL studies on this topic (Chapter 3).

These findings support the modified version of the population mixing hypothesis tested in this research: that a large increase in population mixing over short time periods acts as a trigger for ALL or type 1 diabetes onset. A single ALL/type 1 diabetes causing infection, or an abnormal response to one or more non-specific infections introduced through population mixing, could be the final trigger for either disease. Either scenario would be reliant upon previously low exposure to infection through low population mixing levels in the area. However, this potential explanation is at odds with results of the static population mixing and type 1 diabetes analyses.

High population mixing in early life

There was some evidence to suggest that high population mixing levels encountered in early life were associated with raised incidence of childhood type 1 diabetes in Canterbury. This finding was only based on a few models and was only noted prior to 1992. After 1992 no significant relationships were found between type 1 diabetes and the static population mixing variables. However, the positive findings prior to 1992 were noted in both the 6/7 and the 12/13 year analyses. These results do not agree with the few previous studies on the topic (Cox, 2007, Feltbower et al., 2005, Parslow et al., 2001), nor the hygiene hypothesis which the analyses sought to test. Two main explanations for these findings were proposed. The first is that high population mixing levels in this region (and possibly New Zealand) are detrimental when encountered at any stage in the life-course. It could be that a single type 1 diabetes causing infection is introduced through population mixing, or that type 1 diabetes is triggered by the

stress of one or more non-specific infections to children with weakened immune systems due to low infectious exposure *in utero*. Alternatively, the population mixing measures may not be accurately capturing population mixing exposure in early life for the majority of children as planned.

Wider context

Thus the general picture from these results is that high or increased levels of population mixing are detrimental in terms of the future development of ALL and type 1 diabetes in children. Increased risk of either disease occurred in areas where population mixing had increased the most during 6/7 year time periods. However, the population mixing regression models generally explained only a small proportion of the total geographical variation in each disease. This finding implies that there was residual confounding in the models, and is consistent with the literature which suggests that the causes of both diseases are likely to be multifactorial (Greaves, 2006, Haverkos, 1997). Thus it is possible that individual lifestyle factors associated with improved hygiene standards in affluent countries, working in combination with population mixing in an increasingly mobile world, have important implications for the immune systems and chronic disease profiles of children. Further research would be necessary to assess this hypothesis. Future research directions will be addressed in the final section of this chapter, after a brief overview of the research dissemination strategy.

10.4 Publication strategy

In terms of the dissemination of these research findings, a number of peer reviewed publications are at varying stages of completion. First, a population mixing literature review was published in the *Australasian Epidemiologist* in April 2007 (Miller et al., 2007). This review was considered necessary as many studies have used population mixing measures without much consideration of their theoretical importance and possible limitations. Second, a paper is currently under review for a special edition (International Medical Geography Symposium 2007) of the journal *Social Science and Medicine*, and discusses the results of the population mixing change and type 1 diabetes analyses for the Canterbury region (Miller et al., in review). This research is the first to test the hypothesis that an increase in population mixing over short periods of time could act as the trigger for type 1 diabetes development in children.

The third publication, will consider the small area geographical epidemiology of childhood type 1 diabetes in the Canterbury region. It will focus upon the descriptive differences in the incidence of this disease by different types of area (urban/rural and affluent/deprived), and the geographical patterns of incidence for small areas across the region. Under this latter heading, the results of the spatial-temporal cluster analyses presented in Chapter 8 will also be discussed. While the importance of geography in type 1 diabetes studies has previously been noted (e.g. Brown, 1993, Samuelsson and Lofman, 2004, Staines, 1996), this research is the first to examine the small area geography of type 1 diabetes in a New Zealand setting. A similar paper is also planned to detail the geographical epidemiology of childhood ALL for small areas across the whole of New Zealand.

Fourth, the results of the ALL and population mixing analyses detailed in Chapter 7 will also be written up for publication. Of particular interest to current debates in the field of population mixing and childhood leukaemia, are the generally consistent findings over different lengths of study periods, and the use of a number of different measures of population mixing, and their varying importance. Finally, a paper will be written on the spatial and temporal variations in population mixing for CAUs across New Zealand for a migration and population geography audience. This publication will examine possible explanations for the social gradients noted in many of the population mixing measures.

10.5 Future research

Future lines of research could usefully determine more accurate measures of exposure. Under this general category there are a number of possibilities for improvement. First, the population mixing measures created could be validated using infection data. The current study relies upon the assumption that high in-migration, migrant diversity and general population mobility, result in a high number and range of infections in any given area. In fact, all of the similar studies on population mixing and health conducted to date have been based upon this assumption (e.g. Cox, 2007, Feltbower et al., 2005, Kinlen, 1988, Koushik et al., 2001, Law et al., 2003, Parslow et al., 2001). However, to my knowledge, no study has tested whether this theory is correct, nor examined under which conditions the theory may not be true. To accomplish this would be a difficult task, since not all infections (e.g. common cold) are 'notifiable', and thus numbers are not recorded. However, there are a number of infections which are classed as notifiable diseases within New Zealand, and which can affect children (e.g. mumps, measles, rubella, influenza, rheumatic fever and hepatitis) (Population and Environmental Health Group, 2006). It would be

useful to statistically compare infection rates and the various measures of population mixing in a sample of different types of areas. For example, it would be valuable to assess how well the population mixing measures predict infections in urban compared to rural areas and in areas with high versus low population densities. Comparing the population mixing measures and infections data could help to determine which population mixing measures are optimal for estimating area-level infectious exposure. However, it is questionable as to whether the notifiable disease data available in New Zealand reflect the true incidence of each disease. Differential reporting practices affect the resulting disease rates. Medical practitioners are less likely to be consulted where the illness is not severe enough, where the costs of visiting medical practitioners are too high and where there is a lack of public awareness about the disease. Furthermore, loose case definitions for some diseases and the interest, resources and priorities of local public health services will all impact upon the data collected. In addition, cases are allocated to a geographic locale based on where the case first consulted a medical practitioner, rather than on where they live, or where they caught the infection (Population and Environmental Health Group, 2006).

A second exposure-related advancement would be to establish the optimal geographical scale at which to measure population mixing. The present study has captured population mixing patterns for relatively small statistical areas (CAUs) which contain between 3,000 and 3,500 people. However, smaller units of analysis may be preferable, and data could also be analysed at the meshblock-level. Although this was not possible in the current study, census data are becoming more routinely stored in electronic format for meshblocks. As a result, future studies will be able to test differences between the results at the CAU and meshblock-level. In addition, it is important to consider over what geographical scales and spaces children are normally active. This study (and previous work) assumes that children will (only) be exposed to population mixing in the area immediately around their current residence. Recent papers have critiqued this approach in the field of health geography as a whole (Cummins, 2007, Cummins et al., 2007), although little has been carried out to date on children's usual action spaces, especially within New Zealand. Questionnaire, interview or diary based approaches to document the usual movements of a sample of children might help to shed some light on this topic. Alternatively, obtaining the addresses of the schools attended by children would help to estimate the main areas within which children are active. However, children are likely to be more restricted in their daily movements than adults (O'Brien et al., 2000) and according to the Ministry of Education (2007), most children in New Zealand attend the school closest to where they live. Consequently, the CAU of residence was considered an appropriate choice for this analysis.

A third possible improvement relates to the timing of exposures. Currently improvements in this area are quite difficult to achieve due to our lack of knowledge on the length of time required for early infectious exposure to be protective, and the length of the pre-clinical periods of both diseases. Advances in the medical sphere on both aspects would greatly improve population mixing studies. However, a way around this problem would be to collect address data on the child's whereabouts at birth, especially since early life infectious exposure is likely to impact on immune system development. Population mixing could thus be measured for each child's CAU of residence at birth to ascertain levels of early life infectious exposure. This method has started to be employed in more recent studies of childhood ALL in Europe, with results showing increased risk in areas with high population mixing at birth (Nyari et al., 2006, Rudant et al., 2006). These findings suggest that even very early life exposures to high population mixing are involved in the later development of the disease. Differences between the importance of population mixing at birth and at the time of diagnosis could be assessed in regression models. Furthermore, cluster analyses could be conducted using the CAU of residence at birth and the date of birth as the spatial and temporal locators. Evidence of clustering at birth would suggest that environmental exposures operating either *in utero* or perinatally are important in the instigation of a disease. Address at birth is not currently collected for type 1 diabetes patients in the Canterbury region, but could usefully be incorporated into future data collection practices. Future leukaemia studies could apply to the Registrar General and the relevant ethics committees to have access to the birth records of every ALL case to obtain the street address of residence at birth.

Extending this idea further, address data could be collected at a number of different points within a child's life-course. This enhanced longitudinal approach would allow population mixing levels to be assessed at the place of residence during the mother's pregnancy, at birth, at nursery and school attendance, and at the time of diagnosis. Changes in the levels of population mixing which each child are exposed to could then be assessed in more detail. Population mixing levels could also usefully be measured in the area of nursery and school attendance where this differs to the home address. The importance of the place of nursery and primary school attended has already been noted in cluster analysis studies (Bodington et al., 1995, Heath, 2005). Such an approach would require data collection in the form of questionnaires and interviews with the parents of each child.

Finally, confounding variables which may limit or enhance infectious exposure should be controlled for in future studies. High levels of residential segregation for example could limit the

number of contacts children have with other potential carriers of infection. At the area-level, residential segregation could be calculated for CAUs using the index of dissimilarity (Wong, 2003) to identify areas with high ethnic or social class segregation. However, this would only operate as a proxy for actual social mixing between the different groups. How safe parents perceive an area to be is likely to have consequences for the restrictions they place on children's play and mobility spaces. Data on actual crimes could be incorporated into models as a surrogate measure for crime perceptions. At a smaller spatial scale it would also be interesting to investigate physical barriers to children mixing in their local environment, for example walled and gated housing, lack of cul-de-sacs or long busy roads. Furthermore, interviews with a sample of parents and children could be conducted to directly determine the number, intensity and variety of early social contacts of children from different backgrounds.

A number of other potential risk factors should also be controlled for at the individual-level. For example, the individual ethnicity of cases and their family histories of similar diseases would allow genetic risk factors to be better accounted for. Both diseases are more common in children of European descent and in children whose close relatives also have these diseases. Ideally, information on the number of siblings and birth order of the case child, the number of infections in early life, and attendance at day-care and playgroup facilities should also be controlled for. To obtain such data would require retrospective collection from the parents of cases (and control children) in the form of interviews or questionnaires. This process would be costly and time consuming, and would require further ethical approval. If achieved, these data could be included in multi-level models to assess whether population mixing has an independent effect on these diseases at the area-level.

10.6 Conclusion

The potential health effects of population mixing are difficult to research due to the number of unknown factors involved in the population mixing hypothesis. For example, it is still unproven that infection is the main cause of either ALL or type 1 diabetes in children. As yet, no one infection has been identified as the putative causal factor for either disease. As suggested by Greaves (1988) with regards to childhood ALL, the disease could be caused by a lack of infectious exposure in early life *per se*. Kolb (1994) has more recently extended this argument to type 1 diabetes pathogenesis. Furthermore, and as highlighted by this research, *time* remains an important unknown in this area of research. More specifically; at what age and for how long is exposure to population mixing important? In terms of measuring population mixing exposures

and subsequent disease occurrence, more accurate information regarding lag times between disease initiation and disease diagnosis is also necessary. The majority of these 'unknowns' rely upon further work in the medical sphere. However, this research has added to a large body of epidemiological research supporting a role for infections in the aetiology of childhood ALL and type 1 diabetes. Moreover, this research has specifically identified a role for population mixing in the pathogenesis of these diseases in New Zealand. While these conclusions require further investigation, any clues towards the aetiology of these diseases are worthwhile due to their considerable morbidity and mortality outcomes. Of the 781 children in this study who were diagnosed with ALL between 1980 and 2004, 210 (27 percent) died as a result of this disease.

References

ABBOTT, M. W., WONG, S., GILES, L. C., WONG, S., YOUNG, W. & AU, M. (2003) Depression in older Chinese migrants to Auckland. *Australian and New Zealand Journal of Psychiatry,* 37**,** 445-451.

ABBOTT, M. W., WONG, S., WILLIAMS, M., AU, M. & YOUNG, W. (1999) Chinese migrants' mental health and adjustment to life in New Zealand. *Australian and New Zealand Journal of Psychiatry,* 33**,** 13-21.

ACHENBACH, P., BONIFACIO, E., KOCZWARA, K. & ZIEGLER, A. G. (2005) Natural history of type 1 diabetes. *Diabetes,* 54**,** S25-31.

ADAMSON, P., LAW, G. & ROMAN, E. (2005) Assessment of trends in childhood cancer incidence. *The Lancet,* 365**,** 753.

ADELMAN, A. S., GROVES, F. D., ROURKE, K. O., SINHA, D., HULSEY, T. C., LAWSON, A. B., WARTENBERG, D. & HOEL, D. G. (2007) Residential mobility and risk of childhood acute lymphoblastic leukaemia: an ecological study. *The British Journal of Cancer,* 97**,** 140.

ADELMAN, A. S., MCLAUGHLIN, C. C., WU, X.-C., CHEN, V. W. & GROVES, F. D. (2005) Urbanisation and incidence of acute lymphocytic leukaemia among United States children aged 0-4. *British Journal of Cancer,* 92**,** 2084-2088.

ADVISORY GROUP ON NON-IONISING RADIATION (2001) ELF Electromagnetic fields and the risk of caner. Report of an advisory Group on non-ionising radiation. Documents of the NRPB. Volume 12, no. 1. Chilton, Didcot, National Radiological Protection Board.

AFIFI, A. A., KOTLERMAN, J. B., ETTNER, S. L. & COWAN, M. (2007) Methods for improving regression analysis for skewed continuous or counted responses. *Annual Review of Public Health,* 28**,** 1–10.

AGRESTI, A. (1996) *Categorical data analysis,* New York, Wiley-Interscience.

AHLBOM, A., DAY, N., FEYCHTING, M., ROMAN, E. & *AL., E.* (2000) A pooled analysis of magnetic fields and childhood leukaemia. *The British Journal of Cancer,* 83**,** 692.

AJIKI, W., HANAI, A., TSUKUMA, H., HIYAMA, T. & FUJIMOTO, L. (1994) Incidence of childhood cancer in Osaka, Japan, 1971-1988: Reclassification of registered cases by Birch's scheme using information on clinical diagnosis, histology and primary site. *Japanese Journal of Cancer Research* 85**,** 139-146.

AKERBLOM, H. K. & KNIP, M. (1998) Putative environmental factors in type 1 diabetes. *Diabetes/Metabolim Reviews,* 14**,** 31-67.

AKERBLOM, H. K., VAARALA, O., HYOTY, H., ILONEN, J. & KNIP, M. (2002) Environmental factors in the etiology of type 1 diabetes. *American Journal of Medical Genetics,* 115**,** 18-29.

AKERBLOM, H. K., VIRTANEN, S. M., HAMALAINEN, A., ILONEN, J., SAVILAHTI, E., VAARALA, O., REUNANNEN, A., TERAMO, K. & KNIP, M. (1999) Emergence of diabetes associated autoantibodies in the nutritional prevention of IDDM (TRIGR) project. (Abstract). *Diabetes,* 48**,** SA45.

ALDOUS, J., BARDSLEY, M., DANIELL, R., GAIR, R., JACOBSEN, B., LOWDELL, C., MORGAN, D., STORKLEY, M. & TAYLOR, G. (1999) *Refugee health in London : key issues for public health,* London, Health of Londoners Project.

ALEXANDER, F. E., BOYLE, P., CARLI, P. M., COEBERGH, J. W., DRAPER, G. J., EKBOM, A., LEVI, F., MCKINNEY, P. A., MCWHIRTER, W., MAGNANI, C., MICHAELIS, J., OLSEN, J. H., PERIS-BONET, R., PETRIDOU, E., PUKKALA, E. & VATTEN, L. (1998) Spatial temporal patterns in childhood leukaemia: further evidence for an infectious origin. *British Journal of Cancer,* 77**,** 812-817.

ALEXANDER, F. E., CHAN, L. C., LAM, T. H., YUEN, P., LEUNG, N. K., HA, S. Y., YUEN, H. L., LI, C. K., LI, C. K., LAU, Y. L. & GREAVES, M. F. (1997) Clustering of childhood leukaemia in Hong Kong: association with the childhood peak of common acute lymphoblastic leukaemia and with population mixing. *British Journal of Cancer,* 75**,** 457-463.

ALEXANDER, F. E., LEON, D. A. & CARTWRIGHT, R. A. (1996) Isolation, car ownership, and small area variation in incidence of acute lymphoblastic leukaemia in children. *Paediatric and Perinatal Epidemiology,* 10**,** 411-417.

ALEXANDER, F. E., MCKINNEY, P. A., RICKETTS, T. J. & CARTWRIGHT, R. (1990) Community lifestyle characteristics and risk of acute lymphoblastic leukaemia in children. *The Lancet,* 336**,** 1461-1465.

ALLISON, A. J. & CLARKE, A. J. (2006) Letter to the editor: Further research for consideration in 'the A2 milk case'. *European Journal of Clinical Nutrition,* 60**,** 921-925.

ALTINKAYNAK, S., SELIMOGLU, M. A., TURGUT, A., KILICASLAN, B. & ERTEKIN, V. (2006) Breast-feeding duration and childhood acute leukemia and lymphomas in a sample of Turkish children *Journal of Pediatric Gastroenterology and Nutrition,* 42**,** 568-572.

ALVORD, E. C., JAHNKE, U., FISCHER, E., KIES, M. & DRISCOLL, B. (1987) The multiple causes of multiple sclerosis: the importance of age of infections in childhood. *Journal of Child Neurology,* 2**,** 313-21.

ANDERSON, A. (2006a) Retrievable time: prehistoric colonisation of South Polynesia from the outside in and the inside out. IN BALLANTYNE, T. & MOLOUGHNEY, B. (Eds.) *Disputed histories: imagining New Zealand's past.* Dunedin, Otago University Press.

ANDERSON, L. M. (2006b) Environmental genotoxicants/carcinogens and childhood cancer: bridgeable gaps in scientific knowledge. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis,* 608**,** 136-156.

ANDERSON, L. M., DIWAN, B. A., FEAR, N. T. & ROMAN, E. (2000) Critical windows of exposure for children's health: cancer in human epidemiological studies and neoplasms in experimental animal models. *Environmental Health Perspectives,* 108**,** 573-594.

ANDERSON, R. M. (1982) Directly transmitted viral and bacterial infections of man. IN ANDERSON, R. M. (Ed.) *The Population Dynamics of Infectious Diseases: Theory and Applications.* London, Chapman and Hall.

ANDERSON, R. M. & MAY, R. M. (1979) Population biology of infectious diseases; Part 1. *Nature,* 280**,** 361-367.

ANDERSON, R. M. & MAY, R. M. (1982) Directly transmitted infectious diseases: control by vaccination. *Science,* 215**,** 1053-1060.

ANON (1990) Childhood leukaemia: an infectious disease? *The Lancet,* 336**,** 1477-1479.

ANTECOL, H. & BEDARD, K. (2006) Unhealthy assimilation: why do immigrants converge to American health status levels? *Demography,* 43**,** 337-360.

ARMITAGE, P. & BERRY, G. (1994) *Statistical methods in medical research,* Oxford, Blackwell Scientific Publications.

ATKINSON, M. A., BOWMAN, M. A., CAMPBELL, L., DARROW, B. L., KAUFMAN, D. L. & MACLAREN, N. K. (1994) Cellular immunity to a determinant common to glutamate decarboxylase and coxsackie virus in insulin-dependent diabetes. *Journal Of Clinical Investigation,* 94**,** 2125–2129.

AZAR, S. T., TAMIM, H., BEYHUM, H. N., HABBAL, M. Z. & ALMAWI, W. Y. (1999) Type I (insulin-dependent) diabetes is a Th1- and Th2-mediated autoimmune disease. *Clinical and Diagnostic Laboratory Immunology,* 6**,** 306-310.

BACH, J.-F. (2001) Protective role of infections and vaccinations on autoimmune diseases. *Journal of Autoimmunity,* 16**,** 347-353.

BACH, J. F. (2002) The effect of infections on susceptibility to autoimmune and allergic diseases. *The New England Journal of Medicine,* 347**,** 911.

BACH, J. F. (2005a) Infections and autoimmune diseases. *Journal of Autoimmunity,* 25**,** 74-80.

BACH, J. F. (2005b) Six questions about the hygiene hypothesis. *Cellular Immunology,* 233**,** 158-161.

BACHE, I., BOCK, T., VOLUND, A. & BUSCHARD, K. (1999) Previous maternal abortion, longer gestation, and younger maternal age decrease the risk of type 1 diabetes among male offspring. *Diabetes Care,* 22**,** 1063.

BAILEY, A. J. (2001) Turning transnational: notes on the theorisation of international migration. *International Journal of Population Geography,* 7**,** 413-428.

BAILEY, A. J., WRIGHT, R. A., MOUNTZ, A. & MIYARES, I. M. (2002) (Re)producing Salvadoran transnational geographies. *Annals of the Association of American Geographers,* 92**,** 125-144.

BALARAJAN, R. & BULUSU, L. (1990) Mortality among immigrants in England and Wales, 1979-83. IN BRITTON, M. (Ed.) *Mortality and geography: a review in the mid-1980s England and Wales.* London, Stationary Office.

BARCHAM, M. (1998) The challenge of urban Māori: Reconciling conceptions of indigeneity and social change. *Asia Pacific Viewpoint,* 39**,** 303-314.

BARNETT, R. (2000) Does place of residence matter? Contextual effects and smoking in Christchurch. *New Zealand Medical Journal,* 113**,** 433-5.

BAUDER, H. (2003) Brain abuse or the devaluation of immigrant labour in Canada. *Antipode,* 35**,** 699-717.

BAUMER, J. H., HUNT, L. P. & SHIELD, J. P. H. (1998) Social disadvantage, family composition, and diabetes mellitus: prevalence and outcome. *Archives of Disease in Childhood,* 79**,** 427-430.

BEAGLEHOLE, R., EYLES, E. & PRIOR, I. (1979) Blood pressure and migration in children. *International Journal of Epidemiology,* 8**,** 5-10.

BEDFORD, R., BEDFORD, C., HO, E. & LIDGARD, J. (2002a) The globalisation of international migration in New Zealand: contribution to a debate. *New Zealand Population Review,* 28**,** 69-97.

BEDFORD, R., HO, E. & HUGO, G. (2003) Trans-Tasman migration in context: recent flows of New Zealanders revisited. *People and Place,* 11**,** 53-62.

BEDFORD, R., HO, E. & LIDGARD, J. (2001) Immigration policy and New Zealand's development into the 21 st Century: review and speculation. *Asian and Pacific Migration Journal,* 10**,** 585-616.

BEDFORD, R., HO, E. & LIDGARD, J. (2005) From targets to outcomes: immigration policy in New Zealand, 1996-2003. IN TRLIN, A. D., SPOONLEY, P. & WATTS, N. (Eds.) *New Zealand and international migration. A digest and bibliography No. 4*. . Palmerston North, Department of Sociology, Social Policy and Social Work, Massey University.

BEDFORD, R. D., HO, E. & LIDGARD, J. (2002b) International migration in New Zealand: context, components and policy issues. IN CARMICHAEL, G. A. & DHARMALINGAM, A. (Eds.) *Populations of New Zealand and Australia at the Millenium: A joint special issue of the Journal of Population Research and the New Zealand Population Review.* Canberra and Wellington, Australian Population Association and Population Association of New Zealand.

BELLEC, S., BACCAINI, B., GOUBIN, A., RUDANT, J., RIPERT, M., HEMON, D. & CLAVEL, J. (2008) Childhood leukaemia and population movements in France, 1990-2003. *British Journal of Cancer,* 98**,** 225-231.

BELSON, M., KINGSLEY, B. & HOLMES, A. (2007) Risk factors for acute leukemia in children: a review. *Environmental Health Perspectives,* 115**,** 138-145.

BENER, A., DENIC, S. & GALADARI, S. (2001) Longer breast-feeding and protection against childhood leukaemia and lymphomas. *European Journal of Cancer,* 37**,** 234-238.

BENNETT, D. (2005) Replacing positivism in medical geography. *Social Science & Medicine,* 60**,** 2685-2695.

BENTHAM, G. (1988) Migration and morbidity: implications for geographical studies of disease. *Social Science & Medicine,* 26**,** 49-54.

BENTHAM, G. (1994) Population mixing and sudden infant death syndrome in England and Wales. *International Journal of Epidemiology,* 23**,** 540-544.

BERKELEY, J. & LUNT, H. (2006) Diabetes epidemiology in New Zealand—does the whole picture differ from the sum of its parts? *New Zealand Medical Journal,* 119**,** 1-3.

BETTS, P., MULLIGAN, J., WARD, P., SMITH, B. & WILKIN, T. J. (2005) Increasing body weight predicts the earlier onset of insulin-dependant diabetes in childhood: testing the 'accelerator hypothesis' (2). *Diabetic Medicine,* 22**,** 144-151.

BHATIA, S. & ROBISON, L. L. (1999) Epidemiology of leukemia and lymphoma. *Current Opinion in Hematology,* 6**,** 201.

BINGLEY, P. J., DOUEK, I. F., ROGERS, C. A. & GALE, E. A. M. (2000) Influence of maternal age at delivery and birth order on risk of type 1 diabetes in childhood: prospective population based family study. *British Medical Journal,* 321**,** 420-424.

BIRCH, J. M., ALEXANDER, F. E., BLAIR, V., EDEN, T. O. B., TAYLOR, G. M. & MCNALLY, R. J. Q. (2000) Space-time clustering patterns in childhood leukaemia support a role for infection. *British Journal of Cancer,* 82**,** 1571-1576.

BLACK, D. & INDEPENDENT ADVISORY GROUP (1984) Investigation into the possible increased incidence of cancer in West Cumbria. Report of the Independent Advisory Group. Chairman: Sir Douglas Black. London, HM Stationery Office.

BLAKELY, T., KIRO, C. & WOODWARD, A. (2002) Unlocking the numerator-denominator bias. II: adjustments to mortality rates by ethnicity and deprivation during 1991-94. The New Zealand Census-Mortality Study. *New Zealand Medical Journal,* 115**,** 43-+.

BLOM, L., DAHLQUIST, G., NYSTROM, L., SANDSTROM, A. & WALL, S. (1989) The Swedish diabetes study - social and perinatal determinants for diabetes in childhood. *Diabetologia,* 32**,** 7-13.

BLOM, L., PERSSON, L. A. & DAHLQUIST, G. (1992) A high linear growth is associated with an increased risk of childhood diabetes mellitus. *Diabetologia,* 35**,** 528-533.

BLOMQVIST, M., JUHELA, S., ERKITTÄ, S. & AL., E. (2002) Rotavirus infections and development of diabetes associated antibodies during the first two years of life. *Clinical and Experimental Immunology,* 128 511–515.

BODANSKY, H. J., STAINES, A., STEPHENSON, C., HAIGH, D. & CARTWRIGHT, R. (1992) Evidence for an environmental effect in the aetiology of insulin dependent diabetes in a transmigratory population. *British Medical Journal,* 304**,** 1020-1022.

BODINGTON, M. J., MUZULU, S. I. & BURDEN, A. C. (1995) Spatial clustering in childhood diabetes: evidence of an environmental cause. *Diabetic Medicine,* 12**,** 865-7.

BORCH-JOHNSEN, K., JONER, G., MANDRUP-OULSEN, T., CHRISTY, M., ZACHAU-CHRISTIANSEN, B., KASTRUP, K. & NERUP, J. (1984) Relationship between breast-feeding and incidence rates of insulin-dependent diabetes mellitus. A hypothesis. *The Lancet,* ii**,** 1083-1086.

BORMAN, E. (2004) Health tourism. *British Medical Journal,* 328**,** 60-61.

BORUGIAN, M. J., SPINELLI, J. J., MEZEI, G., WILKINS, R., ABANTO, Z. & MCBRIDE, M. L. (2005) Childhood leukemia and socioeconomic status in Canada. *Epidemiology,* 16**,** 526-531.

BOUTOU, O., GUIZARD, A. V., SLAMA, R., POTTIER, D. & SPIRA, A. (2002) Population mixing and leukaemia in young people around the La Hague nuclear waste reprocessing plant. *British Journal of Cancer,* 87**,** 740-745.

BOYLE, P. (2004) Population geography: migration and inequalities in mortality and morbidity. *Progress in Human Geography,* 28**,** 767-776.

BOYLE, P. & DUKE-WILLIAMS, O. (2004) Does migration exaggerate the relationship between deprivation and self-reported limiting long-term illness? IN BOYLE, P., CURTIS, S., GRAHAM, E. & MOORE, E. (Eds.) *The Geography of Health Inequalities in the Developed World: Views from Britain and North America.* Aldershot, Ashgate.

BOYLE, P., FLOWERDEW, R. & WILLIAMS, A. (1997) Evaluating the goodness of fit in models of sparse medical data: a simulation approach. *International Journal of Epidemiology,* 26**,** 651-656.

BOYLE, P. J. & FLOWERDEW, R. (1993) Modelling sparse interaction matrices: interward migration in Hereford and Worcester, and the underdispersion problem. *Environment and Planning A,* 25**,** 1201-1209.

BOYLE, P. J., HALFACREE, K. H. & ROBINSON, V. (1998) *Exploring contemporary migration,* London., Longman.

BREKKE, H. K. & LUDVIGSSON, J. (2007) Vitamin D supplementation and diabetes-related autoimmunity in the ABIS study. *Pediatric Diabetes,* 8**,** 11-14.

BRENNER, D. J., DOLL, R., GOODHEAD, D. T., HALL, E. J., LAND, C. E., LITTLE, J. B., LUBIN, J. H., PRESTON, D. L., PRESTON, R. J., PUSKIN, J. S., RON, E., SACHS, R. K., SAMET, J. M., SETLOW, R. B. & ZAIDER, M. (2003) Cancer risks attributable to low doses of ionizing radiation: Assessing what we really know. *Proceedings of the National Academy of Sciences,* 100**,** 13761-13766.

BRESLOW, N. E. & DAY, N. E. (1980) *Statistical methods in cancer research: Vol. 1. The analysis of case-control studies,* Lyon, International Agency for Research on Cancer (IARC).

BRIMBLECOMBE, N., DORLING, D. & SHAW, M. (2000) Migration and geographical inequalities in health in Britain. *Social Science & Medicine,* 50**,** 861-878.

BROWN, L. J. (1993) Genetics and the environment: understanding geographical variations in the incidence of childhood diabetes. *New Zealand Geographer,* 49**,** 32-39.

BROWN, L. J. & SCOTT, R. S. (1988) A population based diabetes register-development and applications. *Community Health Studies,* 12**,** 437-443.

BROWN, M. (1995) Ironies of distance: an ongoing critique of the geographies of AIDS. *Environment and Planning D,* 13**,** 159-183.

BROWN, T. & DUNCAN, C. (2000) London's burning: recovering other geographies of health. *Health & Place,* 6**,** 363-375.

BRUINING, G. J. (2000) Association between infant growth before onset of juvenile type-1 diabetes and autoantibodies to IA-2. *The Lancet,* 356**,** 655-656.

BUCHANAN, N. (2004) The effect of urban growth on Christchurch travel patterns. *Geography.* Christchurch, University of Canterbury.

BULLIARD, J. L. & REEDER, A. (2001) Getting the message across: sun protection information in media weather reports in New Zealand. *New Zealand Medical Journal,* 114**,** 67-70.

BYERS, A. L., ALLORE, H., GILL, T. M. & PEDUZZI, P. N. (2003) Application of negative binomial modeling for discrete outcomes: a case study in aging research. *Journal of Clinical Epidemiology,* 56**,** 559-564.

CALDER, J. & SAPSFORD, R. (2006) Statistical techniques. IN SAPSFORD, R. & JUPP, V. (Eds.) *Data collection and analysis.* London Sage Publications.

CAMERON, A. C. & TRIVEDI, P. K. (1998) *Regression analysis of count data,* Cambridge, Cambridge University Press.

CAMPBELL-STOKES, P. L. & TAYLOR, B. J. (2005) Prospective incidence study of diabetes mellitus in New Zealand children aged 0 to 14 years. *Diabetologia,* 48**,** 643-648.

CARBALLO, M., DIVINO, J. J. & ZERIC, D. (1998) Migration and health in the European Union. *Tropical Medicine and International Health,* 3**,** 936-944.

CARDWELL, C. R., CARSON, D. J. & PATTERSON, C. C. (2005) Parental age at delivery, birth order, birth weight and gestational age are associated with the risk of childhood Type 1 diabetes: a UK regional retrospective cohort study. *Diabetic Medicine,* 22**,** 200-206.

CARDWELL, C. R., CARSON, D. J. & PATTERSON, C. C. (2006) Higher incidence of childhood-onset type 1 diabetes mellitus in remote areas: a UK regional small-area analysis. *Diabetologia,* 49**,** 2074-2077.

CARDWELL, C. R., CARSON, D. J. & PATTERSON, C. C. (2007) Secular trends, disease maps and ecological analyses of the incidence of childhood onset Type 1 diabetes in Northern Ireland, 1989-2003. *Diabetic Medicine,* 24**,** 289-295.

CASTLES, S. (2000) International migration at the beginning of the twenty-first century: global trends and issues. *International Social Science Journal,* 165**,** 269-281.

CASTLES, S. & MILLER, M. J. (1998) *The age of migration: international population movements in the modern world,* New York, Guilford Press.

CASU, A., CARLINI, M., CONTU, A. & *AL., E.* (2000) Type 1 diabetes in Sardinia is not linked to nitrate levels in drinking water. *Diabetes Care,* 23**,** 1043-1044.

CHAMPSAUR, H., BOTTAZZO, G., BERTRAMS, J., ASSAN, R. & BACH, G. (1982) Virologic, immunologic and genetic factors in insulin-dependent diabetes mellitus. *Journal Of Pediatrics,* 100**,** 15-20.

CHAN, L. C., LAM, T. H., LI, C. K., LAU, Y. L., LI, C. K., YUEN, H. L., LEE, C. W., HA, S. Y., YUEN, P. M., LEUNG, N. K., PATHEAL, S. L., GREAVES, M. F. & ALEXANDER, F. E. (2002) Is the timing of exposure to infection a major determinant of acute lymphoblastic leukaemia in Hong Kong? *Paediatric and Perinatal Epidemiology,* 16**,** 154-165.

CHERUBINI, V., CARLE, F., GESUITA, R. & *AL., E.* (1999) Large incidence variation of type 1 diabetes in central-southern Italy 1990-1995: lower risk in rural areas. *Diabetologia,* 42**,** 789-792.

CHEUNG, P. & SPEARS, G. (1995) Psychiatric morbidity among New-Zealand Cambodians - the role of psychosocial factors. *Social Psychiatry and Psychiatric Epidemiology,* 30**,** 92-97.

CHIARELLI, F., DAHL-JØRGENSEN, K. & KIESS, W. (Eds.) (2005) *Diabetes in childhood and adolescence,* New York Karger.

CHINN, S., HUGHES, J. M. & RONA, R. J. (1998) Trends in growth and obesity in ethnic groups in Britain. *Arch Dis Child,* 78**,** 513-517.

CLARK, B. R., FERKETICH, A. K., FISHER, J. L., RUYMANN, F. B., HARRIS, R. E. & WILKINS, J. R. (2007) Evidence of population mixing based on the geographical distribution of childhood leukemia in Ohio. *Pediatric Blood & Cancer,* 49**,** 797-802.

CLARKSON, T. S., MARTIN, R. J., RUDOLPH, J. & GRAHAM, B. W. L. (1996) Benzene and toluene in New Zealand air. *Atmospheric Environment,* 30**,** 569-577.

CLASSEN, J. B. (1996) Diabetes epidemic follows hepatitus B immunisation program. *New Zealand Medical Journal,* 109**,** 195.

CLIFF, A. D., HAGGETT, P. & ORD, J. K. (1986) *Spatial aspects of influenza epidemics,* London, Pion.

COMMITTEE ON MEDICAL ASPECTS OF RADIATION IN THE ENVIRONMENT (1986) First report: the implications of the new data on the releases from Sellafield in the 1950s for the conclusions of the Report on the Investigation of the Possible Increased Incidence of Cancer in West Cumbria. Chairman Bobrow, M. London, HM Stationery Office.

COMMITTEE ON MEDICAL ASPECTS OF RADIATION IN THE ENVIRONMENT (1988) Second report: investigation of the possible increased incidence of leukaemia in young people near the Dounreay Nuclear Establishment, Caithness, Scotland. Chairman Bobrow, M. London, HM Stationery Office.

COMMITTEE ON MEDICAL ASPECTS OF RADIATION IN THE ENVIRONMENT (1996) Fourth Report. The incidence of cancer and leukaemia in young people in the vicinity of the Sellafield site, West Cumbria: Further studies and an update of the situation since the publication of the report of the Black Advisory Group in 1984. . London, Department of Health.

COMMITTEE ON MEDICAL ASPECTS OF RADIATION IN THE ENVIRONMENT (2002) Seventh Report. Parents occupationally exposed to radiation prior to the conception of their children: A review of the evidence concerning the incidence of cancer in their children. . National Radiological Protection Board.

CONRADSON, D. (2005) Landscape, care and the relational self: therapeutic encounters in rural England. *Health & Place,* 11**,** 337-348.

COOKE, J. V. (1942) The incidence of acute leukemia in children. *Journal of the American Medical Association,* 119**,** 547-550.

COOKE, K. R., GRAY, A. J., BURRY, A. F. & STEWART, R. J. (1988) Cancer Registration in New Zealand: report of the Cancer Registration Working Group. Wellington, Department of Statistics.

COUPER, J. (2007) Type 1 and Type 2 diabetes - overlap or overlay? *New Zealand Paediatric Society.* Canterbury University, New Zealand.

COUPER, J. J., STEELE, C., BERESFORD, S., POWELL, T., MCCAUL, K., POLLARD, A., GELLERT, S., TAIT, B., HARRISON, L. C. & COLMAN, P. G. (1999) Lack of association between duration of breast-feeding or introduction of cow's milk and development of islet autoimmunity. *Diabetes,* 48**,** 2145.

COX, M. (2007) Geographic variations in the incidence of diabetes mellitus in Tayside, Scotland. *Geography Department.* University of St Andrews.

CRAMPTON, P., SALMOND, C. & SUTTON, F. (1997) Research Report No 5: NZDep91 Index of Deprivation. Wellington, Health Services Research Centre.

CROSIGNANI, P., TITTARELLI, A., BORGINI, A., CODAZZI, T., ROVELLI, A., PORRO, E., CONTIERO, P., BIANCHI, N., TAGLIABUE, G., FISSI, R., ROSSITTO, F. & BERRINO, F. (2004) Childhood leukemia and road traffic: A population-based case-control study. *International Journal of Cancer,* 108**,** 596-599.

CROW, Y. J., ALBERTI, K. G. & PARKIN, J. M. (1991) Insulin dependent diabetes in childhood and material deprivation in northern England, 1977-86. *British Medical Journal* 303**,** 158-160.

CRUMP, J. A., MURDOCH, D. R. & BAKER, M. G. (2001) Emerging infectious diseases in an island ecosystem: the New Zealand perspective. *Emerging Infectious Diseases,* 7**,** 767-772.

CUMMINS, S. (2007) Commentary: Investigating neighbourhood effects on health - avoiding the 'Local Trap'. *International Journal of Epidemiology,* 36**,** 355-357.

CUMMINS, S., CURTIS, S., DIEZ-ROUX, A. V. & MACINTYRE, S. (2007) Understanding and representing 'place' in health research: a relational approach. *Social Science & Medicine,* 65**,** 1825-1838

CURTIS, S. & TAKET, A. (1996) *Health and societies: changing perspectives,* London, Edward Arnold.

DABELEA, D., D'AGOSTINO, R. B., JR., MAYER-DAVIS, E. J., PETTITT, D. J., IMPERATORE, G., DOLAN, L. M., PIHOKER, C., HILLIER, T. A., MARCOVINA, S. M., LINDER, B., RUGGIERO, A. M., HAMMAN, R. F. & FOR THE, S. F. D. I. Y. S. G. (2006) Testing the accelerator hypothesis: body size, beta-cell function, and age at onset of type 1 (autoimmune) diabetes. *Diabetes Care,* 29**,** 290-294.

DACOU-VOUTETAKIS, C., KARAVANAKI, K. & TSOKA-GENNATAS, H. (1995) National data on the epidemiology of IDDM in Greece. Cases diagnosed in 1992. *Diabetes Care,* 18**,** 552-554.

DAHLQUIST, G., BLOM, L., PERSSON, L.-A., SANDSTROM, A. I. M. & WALL, S. G. I. (1990) Dietary factors and the risk of developing insulin dependent diabetes in childhood. *British Medical Journal,* 300**,** 1302–1306.

DAHLQUIST, G. & KALLEN, B. A. J. (1996) Time-space clustering of date at birth in childhood onset diabetes. *Diabetes Care,* 19**,** 328-332.

DANEMAN, D. (2005) Is the 'Accelerator Hypothesis' worthy of our attention? *Diabetic Medicine,* 22**,** 115-117.

DANIELS, J. L., OLSHAN, A. F. & SAVITZ, D. A. (1997) Pesticides and childhood cancers. *Environmental Health Perspectives,* 105**,** 1068-1077.

DARLOW, B. A., WILLIS, J. A., SCOTT, R. S. & *AL., E.* (2004) Type 1 diabetes in New Zealand children and adolescents: Incidence, prevalence and mortality. *Pediatric Research,* 56**,** 474-474.

DEAN, G. (1971) Multiple Sclerosis in migrants to South Africa. *Israel Journal of Medical Sciences,* 7.

DESANDES, E., CLAVEL, J., BERGER, C., BERNARD, J.-L., BLOUIN, P., DE LUMLEY, L., DEMEOCQ, F., FREYCON, F., GEMBARA, P., GOUBIN, A., LE GALL, E., PILLON, P., SOMMELET, D., TRON, I. & LACOUR, B. (2004) Cancer incidence among children in France, 1990-1999. *Pediatric Blood & Cancer,* 43**,** 749-757.

DESOUZA, R. (2005) Transforming possibilities of care: Goan migrant motherhood in New Zealand. *Contemporary nurse : a journal for the Australian nursing profession.,* 20**,** 87-101.

DIABETES EPIDEMIOLOGY RESEARCH INTERNATIONAL GROUP, D. (1987) Preventing insulin-dependent diabetes mellitus: the environmental challenge. *British Medical Journal,* 295**,** 479-481.

DIAMOND (2006) Incidence and trends of childhood type 1 diabetes worldwide 1990–1999. *Diabetic Medicine,* 23**,** 857–866.

DICKINSON, H. O. (2005) The causes of childhood leukaemia. *BMJ,* 330**,** 1279-1280.

DICKINSON, H. O., HAMMAL, D. M., BITHELL, J. F. & PARKER, L. (2002) Population mixing and childhood leukaemia and non-Hodgkin's lymphoma in census wards in England and Wales, 1966-87. *British Journal of Cancer,* 86**,** 1411-1413.

DICKINSON, H. O. & PARKER, L. (1999) Quantifying the effect of population mixing on childhood leukaemia risk: the Seascale cluster. *British Journal of Cancer,* 81**,** 144-151.

DICKINSON, H. O. & PARKER, L. (2002) Leukaemia and non-Hodgkin's lymphoma in children of Sellafield male radiation workers. *International Journal of Cancer,* 101**,** 100.

DOCKERTY, J. D., BECROFT, D. M. O., LEWIS, M. E. & WILLIAMS, S. M. (1997) The accuracy and completeness of childhood cancer registration in New Zealand. *Cancer Causes and Control* 8**,** 857-864.

DOCKERTY, J. D., COX, B., BORMAN, B. & SHARPLES, K. J. (1996a) Population mixing and the incidence of childhood leukaemias: retrospective comparison in rural areas of New Zealand. *British Medical Journal,* 312**,** 1203-1204.

DOCKERTY, J. D., COX, B. & COCKBURN, M. G. (1996b) Childhood leukaemias in New Zealand: time trends and ethnic differences. *British Journal of Cancer,* 73**,** 1141-1147.

DOCKERTY, J. D., DRAPER, G., VINCENT, T., ROWAN, S. D. & BUNCH, K. J. (2001) A case-control study of parental age, parity, and socioeconomic level in relation to childhood cancers. *International Journal of Epidemiology,* 30**,** 1428-1437.

DOCKERTY, J. D., ELWOOD, J. M., SKEGG, D. C. & HERBISON, G. P. (1998) Electromagnetic field exposures and childhood cancers in New Zealand. *Cancer Causes and Control,* 9**,** 299-309.

DOCKERTY, J. D., ELWOOD, J. M., SKEGG, D. C. & HERBISON, G. P. (1999a) Correction: Electromagnetic field exposures and childhood cancers in New Zealand *Cancer Causes and Control* 1998; 9(3): 299-309. *Cancer Causes and Control,* 10**,** 641.

DOCKERTY, J. D., ELWOOD, J. M., SKEGG, D. C. G. & HERBISON, G. P. (1999b) Electromagnetic field exposures and childhood leukaemia in New Zealand. *The Lancet,* Vol 354 1967-1968.

DOCKERTY, J. D., SHARPLES, K. J. & BORMAN, B. (1999c) An assessment of spatial clustering of leukaemias and lymphomas among young people in New Zealand. *Journal of Epidemiology and Community Health,* 53**,** 154-158.

DOCKERTY, J. D., SKEGG, D. C. G., ELWOOD, J. M., HERBISON, G. P., BECROFT, D. M. O. & LEWIS, M. E. (1999d) Infections, vaccinations, and the risk of childhood leukaemia. *British Journal of Cancer,* 80**,** 1483-1489.

DOLL, R. (1999) The Seascale cluster: a probable explanation. *British Journal of Cancer,* 81**,** 3-5.

DOLL, R., EVANS, H. J. & DARBY, S. C. (1994) Paternal exposure not to blame. *Nature,* 367**,** 678-680.

DOLL, R. & WAKEFORD, R. (1997) Risk of childhood cancer from fetal irradiation. *Br J Radiol,* 70**,** 130-139.

DORN, M. & LAWS, G. (1994) Social theory, body politics, and medical geography: extending Kearns's invitation. *The Professional Geographer,* 46**,** 106-110.

DRAPER, G. J., HEAF, M. M. & KINNIER WILSON, L. M. (1977) Occurrence of childhood cancers among sibs and estimation of familial risks. *Journal of Medical Genetics,* 14**,** 81–90.

DRAPER, G. J., VINCENT, T. J., O'CONNOR, C. M. & STILLER, C. A. (1991) Socio-economic factors and variations in incidence rates between county districts. IN DRAPER, G. J. (Ed.) *The geographical epidemiology of childhood leukaemia and non-Hodgkin lymphomas in Great Britain, 1966–83* London., OPCS.

DU PREL, J. B., ICKS, A., GRABERT, M., HOLL, R. W., GIANI, G. & ROSENBAUER, J. (2007) Socioeconomic conditions and type 1 diabetes in childhood in North Rhine-Westphalia, Germany. *Diabetologia,* 50**,** 720-728.

DUARTE-DAVIDSON, R., COURAGE, C., RUSHTON, L. & LEVY, L. (2001) Benzene in the environment: An assessment of the potential risks to the health of the population. *Occupational and Environmental Medicine,* 58**,** 2.

DUNCAN, C., JONES, K. & MOON, G. (1996) Health-related behaviour in context: a multilevel modelling approach. *Social Science & Medicine,* 42**,** 817-830.

DWYER, C. (1999) Migrations and diasporas. IN CLOKE, P. J., CRANG, P. & GOODWIN, M. (Eds.) *Introducing human geographies.* London, Arnold.

DYCK, I. (1995) Hidden geographies: the changing lifeworlds of women with multiple sclerosis. *Social Science & Medicine,* 40**,** 307-320.

DYCK, I. & DOSSA, P. (2007) Place, health and home: gender and migration in the constitution of healthy space. *Health & Place,* 13**,** 691-701.

DYCK, I. & KEARNS, R. (1995) Transforming the relations of research: towards culturally safe geographies of health and healing. *Health & Place,* 1**,** 137-147.

EARICKSON, R. J. (2007) Introduction to special issue: Eleventh International Medical Geography Symposium. *Social Science & Medicine,* 65**,** 1-6.

EISENBARTH, G. S. (1986) Type I diabetes mellitus. A chronic autoimmune disease. *The New England Journal of Medicine,* 314**,** 1360–68.

ELFORD, J. & BEN-SHLOMO, Y. (2004) Geography and migration with special relevance to cardiovascular disease. IN KUH, D. & BEN-SHLOMO, Y. (Eds.) *A life course approach to chronic disease epidemiology.* 2nd ed. Oxford Oxford University Press.

ELLIOT, R. B. & MARTIN, J. M. (1984) Dietary protein: a trigger of insulin-dependent diabetes in the BB rat? *Diabetologia,* 26**,** 297-9.

ELLIOTT, R. B., HARRIS, D. P., HILL, J. P., BIBBY, N. J. & WASMUTH, H. E. (1999) Type I (insulin-dependent) diabetes mellitus and cow milk: casein variant consumption. *Diabetologia,* 42**,** 292.

ELLIOTT, S. J. & GILLIE, J. (1998) Moving experiences: a qualitative analysis of health and migration. *Health & Place,* 4**,** 327-339.

ESTEVE, J., BENHAMOU, E. & RAYMOND, L. (1994) *Statistical methods in cancer research: Descriptive epidemiology,* Lyon, International Agency for Research on Cancer.

EURODIAB (1998) Familial risk of type 1 diabetes in European children. *Diabetologia,* 41**,** 1151-1156.

EURODIAB (2000) Infections and vaccinations as risk factors for childhood type I (insulin-dependent) diabetes mellitus: a multicentre case-control investigation. *Diabetologia,* 43**,** 47-53.

EURODIAB (2002) Rapid early growth is associated with increased risk of childhood type 1 diabetes in various European populations. *Diabetes Care,* 25**,** 1755(6).

EURODIAB SUBSTUDY 2 STUDY GROUP (1999) Vitamin D supplement in early childhood and risk for Type I (insulin-dependent) diabetes mellitus. *Diabetologia,* 42**,** 51-54.

EVANS, J. (1987) Introduction: migration and health. *International Migration Review,* 21**,** v-xiv.

EVANS, J. M. M., NEWTON, R. W., RUTA, D. A., MACDONALD, T. M. & MORRIS, A. D. (2000) Socio-economic status, obesity and prevalence of type 1 and type 2 diabetes mellitus. *Diabetic Medicine,* 17**,** 478-480.

FEAR, N. T., MCKINNEY, P. A., PATTERSON, C. C., PARSLOW, R. C. & BODANSKY, H. J. (1999) Childhood type 1 diabetes mellitus and parental occupations involving social mixing and infectious contacts: two population-based case-control studies. *Diabetic Medicine,* 16**,** 1025-1029.

FELTBOWER, R. G., BODANSKY, H. J., MCKINNEY, P. A., HOUGHTON, J., STEPHENSON, C. R. & HAIGH, D. (2002) Trends in the incidence of childhood diabetes in south Asians and other children in Bradford, UK. *Diabetic Medicine,* 19**,** 162-166.

FELTBOWER, R. G., MANDA, S. O. M., GILTHORPE, M. S., GREAVES, M. F., PARSLOW, R. C., KINSEY, S. E., BODANSKY, H. J. & MCKINNEY, P. A. (2005) Detecting small-area similarities in the epidemiology of childhood acute lymphoblastic leukemia and diabetes mellitus, type 1: a bayesian approach. *American Journal of Epidemiology,* 161**,** 1168-1180.

FELTBOWER, R. G., MCKINNEY, P. A., GREAVES, M. F., PARSLOW, R. C. & BODANSKY, H. J. (2004) International parallels in leukaemia and diabetes epidemiology. *Archives of Disease in Childhood,* 89**,** 54-56.

FELTBOWER, R. G., MCNALLY, R. J. Q., PARKER, L., BODANSKY, H. J., CAMPBELL, F. M. & MCKINNEY, P. A. (2006) Further clues to the aetiology of type 1 diabetes: spatial clustering amongst 0-29 year olds in Yorkshire, UK. *Diabetologia,* 49**,** 167-167.

FEUDTNER, C. (2003) *Bittersweet : diabetes, insulin, and the transformation of illness,* Chapel Hill, University of North Carolina Press.

FIELD, A. P. (2005) *Discovering statistics using SPSS: (and sex, drugs and rock 'n' roll),* London, Sage.

FIELDING, T. (1992) Migration and Culture. IN CHAMPION, T. & FIELDING, T. (Eds.) *Migration processes and patterns: volume I. research progress and prospects.* London, Belhaven Press.

FILIPPI, C. & VON HERRATH, M. (2005) How viral infections affect the autoimmune process leading to type 1 diabetes. *Cellular Immunology,* 233**,** 125-132.

FINDLAY, A. & GRAHAM, E. (1991) The challenge facing population geography. *Progress in Human Geography,* 15**,** 149-62.

FINDLAY, A. M. & STOCKDALE, A. (2003) The temporal and social embeddedness of migration: a methodological exploration using biographical analysis. *Geography Research Forum,* 23**,** 4-29.

FINE, P. E. M. (1993) Herd immunity: history, theory and practice. *Epidemiologic Reviews,* 15**,** 265-302.

FLOWERDEW, R., MANLEY, D. J. & SABEL, C. E. (2008) Neighbourhood effects on health: does it matter where you draw the boundaries? *Social Science & Medicine,* 66**,** 1241-1255.

FOLIAKI, S. A., KOKAUA, J., SCHAAF, D. & TUKUITONGA, C. (2006) Twelve-month and lifetime prevalences of mental disorders and treatment contact among Pacific people in Te Rau Hinengaro: The New Zealand Mental Health Survey. *Australian and New Zealand Journal of Psychiatry,* 40**,** 924-934.

FORREST, J. M., MENSER, M. A. & HARLEY, J. D. (1967) A twenty five year follow up of congenital rubella. *The Lancet,* ii**,** 1347–1350.

FOTHERINGHAM, A. S. & WONG, D. W. S. (1991) The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A,* 23**,** 1025-1044.

FOURLANOS, S., NARENDRAN, P., BYRNES, G. B., COLMAN, P. G. & HARRISON, L. C. (2004) Insulin resistance is a risk factor for progression to Type 1 diabetes. *Diabetologia,* 47**,** 1661-1667.

FOX, J. P., ELVEBACK, L., SCOTT, W., GATEWOOD, L. & ACKERMAN, E. (1971) Herd immunity: basic concept and relevance to public health immunization practices. *American Journal of Epidemiology,* 94**,** 179-89.

FRENK, J. & GOMEZ-DANTES, O. (2002) Globalization and the challenges to health systems. *Health Affairs,* 21**,** 160-165.

FRENK, J., SEPULVEDA, J., GOMEZ-DANTES, O., MCGUINNESS, M. J. & KNAUL, F. (1997) The future of world health: the new world order and international health. *British Medical Journal,* 314**,** 1404-.

FRIESEN, W. (2003) Population drain or exchange: internal migration between Auckland and the rest of New Zealand? *New Zealand Journal of Geography,* 115**,** 44-48.

GALE, K. B., FORD, A. M., REPP, R., BORKHARDT, A., KELLER, C., EDEN, O. B. & GREAVES, M. F. (1997) Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots. *Proceedings of the National Academy of Sciences,* 94**,** 13950-4.

GALLO, R. C., ESSEX, M. E. & GROSS, L. (1984) *Human T-cell leukaemia/lymphoma virus,* New York, Cold Spring Harbour Laboratory.

GAMBLE, D. R. & TAYLOR, K. W. (1969) Seasonal incidence of diabetes mellitus. *British Medical Journal,* 3**,** 631–633.

GAMBLE, D. R., TAYLOR, K. W. & CUMMING, H. (1973) Coxsackie viruses and diabetes mellitus. *British Medical Journal,* 4**,** 260-262.

GARDNER, M. J. (1991) Father's occupational exposure to radiation and the raised level of childhood leukemia near the Sellafield nuclear plant. *Environmental Health Perspectives,* 94**,** 5-7.

GARDNER, M. J. (1992) Paternal occupations of children with leukemia. *British Medical Journal,* 305**,** 715.

GARDNER, M. J., SNEE, M. P., HALL, A. J., POWELL, C. A., DOWNES, S. & TERRELL, J. D. (1990) Results of case-control study of leukaemia and lymphoma among young people near Sellafield nuclear plant in West Cumbria. *British Medical Journal,* 300**,** 423–429.

GARDNER, W., MULVEY, E. P. & SHAW, E. C. (1995) Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin,* 118**,** 392-404.

GATRELL, A. C. (2002) *Geographies of health: an introduction,* Oxford, Blackwell.

GELLERT, G. A. (1993) International migration and control of communicable diseases. *Social Science & Medicine,* 37**,** 1489-1499.

GENUTH, S. (2006) Type 1 diabetes mellitus: pathogenesis. ACP Medicine Online.

GERSTEIN, H. C. (1994) Cow's milk exposure and type I diabetes mellitus. A critical overview of the clinical literature. *Diabetes Care,* 17**,** 13-19.

GESLER, W. M. (1992) Therapeutic landscapes: medical issues in light of the new cultural geography. *Social Science & Medicine,* 34**,** 735-746.

GIBBON, C., SMITH, T., EGGER, P., BETTS, P. & PHILLIPS, D. (1997) Early infection and subsequent insulin dependent diabetes. *Archives of Disease in Childhood,* 77**,** 384-85.

GIBBON, C., SMITH, T., EGGER, P., BETTS, P. R. & PHILLIPS, D. I. W. (1995) Protection against Type-1 (Insulin-Dependent) Diabetes by Infection in Infancy - Evidence for the Hygiene Hypothesis. *Diabetologia,* 38**,** A99-A99.

GILES, G. G. (1983) The utility of the relative risk ratio in geographic epidemiology: Hodgkin's disease in Tasmania 1972-1980. IN MCGLASHAN, N. D. & BLUNDEN, J. R. (Eds.) *Geographical Aspects of Health.* London, Academic Press.

GILHAM, C., PETO, J., SIMPSON, J., ROMAN, E., EDEN, T. O. B., GREAVES, M. F. & ALEXANDER, F. E. (2005) Day care in infancy and risk of childhood acute lymphoblastic leukaemia: findings from UK case-control study *British Medical Journal,* 330**,** 1294-1300.

GILLESPIE, K. M. (2006) Type 1 diabetes: pathogenesis and prevention. *CMJA,* 175**,** 165-170.

GILLESPIE, K. M., BAIN, S. C., BARNETT, A. H., BINGLEY, P. J., CHRISTIE, M. R., GILL, G. V. & GALE, E. A. M. (2004) The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. *The Lancet,* 364**,** 1699-1700.

GILLIAM, A. G. & WALTER, W. A. (1958) Trends in mortality from leukemia in the United States 1921-55. *Public Health Reports 73.* Washington.

GILMAN, E. A. & KNOX, E. G. (1998) Geographical distribution of birth places of children with cancer in the UK *British Journal of Cancer,* 77**,** 842-849.

GINSBERG-FELLNER, F., KLEIN, E., DOBERSEN, M. & AL., E. (1980) The interrelationship of congenital rubella (CR) and insulin dependent diabetes mellitus (IDDM). *Pediatric Research,* 14**,** 57.

GLASS, A. G., MANTEL, N., GUNZ, F. W. & *AL., E.* (1971) Time-space clustering of childhood leukemia in New Zealand. *Journal of the National Cancer Institute,* 47**,** 329-36.

GLASS, D., GRAY, C., JOLLEY, D., GIBBONS, C., SIM, M., FRITSCH, I. L., ADAMS, G., BISBY, J. & MANUELL, R. (2003) Leukemia risk associated with low-level benzene exposure. *Epidemiology,* 14**,** 569-77.

GLERUM, M. & ROBINSON, B. H. (1989) Could bovine serum albumin be the initiating antigen ultimately responsible for the development of insulin dependent diabetes mellitus? *Diabetes Research,* 10**,** 103-107.

GOOGLE (2007) Google Maps. Accessed March 13 2007, from: http://maps.google.com/maps. Google.

GORDIS, L. (2000) *Epidemiology,* Philadephia, W.B. Saunders.

GRAHAM, E. (2004) The past, present and future of population geography: reflections on Glenn Trewartha's address fifty years on. *Population, Space and Place,* 10**,** 289-294.

GREAVES, M. (2007) Darwinian medicine: a case for cancer. *Nature Reviews Cancer,* 7**,** 213-221.

GREAVES, M. F. (1988) Speculations on the cause of childhood acute lymphoblastic leukaemia. *Leukemia,* 2**,** 120-125.

GREAVES, M. F. (1997) Aetiology of acute leukaemia. *The Lancet,* 349**,** 344-349.

GREAVES, M. F. (2002) Childhood leukaemia: science medicine and the future. *British Medical Journal,* 324**,** 283-287.

GREAVES, M. F. (2006) Infection, immune responses and the aetiology of childhood leukaemia. *Nature Reviews Cancer,* 6 193-203.

GREAVES, M. F. & ALEXANDER, F. E. (1993) An infectious etiology for common acute lymphoblastic leukemia in childhood? *Leukemia,* 7**,** 349-360.

GREAVES, M. F., MAIA, A. T., WIEMELS, J. L. & FORD, A. M. (2003) Leukemia in twins: lessons in natural history. *Blood,* 102**,** 2321-2333.

GREAVES, M. F. & WIEMELS, J. (2003) Origins of chromosome translocations in childhood leukaemia. *Nature Reviews Cancer,* 3**,** 639.

GREER, R. M., ROGERS, M. A., BOWLING, F. G., BUNTAIN, H. M. & ET AL. (2007) Australian children and adolescents with type 1 diabetes have low vitamin D levels. *Medical Journal of Australia,* 187**,** 59.

GROSSMAN, J. (1995) What's hiding under the sink: dangers of household pesticides. *Environmental Health Perspectives,* 103**,** 550–554.

GUIZARD, A. V., BOUTOU, O., POTTIER, D., TROUSSARD, X., PHEBY, D., LAUNOY, G., SLAMA, R., SPIRA, A. & ARKM (2001) The incidence of childhood leukaemia around the La Hague nuclear waste reprocessing plant (France): a survey for the years 1978-1998. *Journal of Epidemiology and Community Health,* 55**,** 469-474.

GUNZ, F. W. & SPEARS, G. F. (1968) Distribution of acute leukaemia in time and space studies in New Zealand. *British Medical Journal,* 4**,** 604-608.

GUSHULAK, B. D. & MACPHERSON, D. W. (2006) The basic principles of migration health: population mobility and gaps in disease prevalence. *Emerging Themes in Epidemiology,* 3.

GUSTAFSSON, B. & CARSTENSEN, J. (2000) Space-time clustering of childhood lymphatic leukaemias and non-Hodgkin's lymphomas in Sweden. *European Journal of Epidemiology,* 16**,** 1111-1116.

HALFACREE, K. H. (1995) Household migration and the structuration of patriarchy: evidence from the USA. *Progress in Human Geography. ,* 19**,** 159-182.

HALFACREE, K. H. & BOYLE, P. (1993) The challenge facing migration research: the case for a biographical approach. *Progress in Human Geography,* 17**,** 333-348.

HARDING, S., BALARAJAN, R. & BALARAJAN, R. (1996) Patterns of mortality in second generation Irish living in England and Wales: longitudinal study. *British Medical Journal,* 312**,** 1389-1392.

HARRIS, H. F. (1899) A case of diabetes mellitus quickly following mumps. *Boston Medical and Surgical Journal*.

HAVERKOS, H. W. (1997) Could the aetiology of IDDM be multifactorial? *Diabetologia,* 40**,** 1235-1240.

HAYNES, A., BOWER, C., BULSARA, M. K., FINN, J., JONES, T. W. & DAVIS, E. A. (2007a) Perinatal risk factors for childhood type 1 diabetes in Western Australia - a population-based study (1980-2002). *Diabetic Medicine,* 24**,** 564-570.

HAYNES, A., BOWER, C., BULSARA, M. K., JONES, T. W. & DAVIS, E. A. (2004) Continued increase in the incidence of childhood type 1 diabetes in a population-based Australian sample (1985-2002). *Diabetologia,* 47**,** 866-870.

HAYNES, A., BULSARA, M. K., BOWER, C., CODDE, J. P., JONES, T. W. & DAVIS, E. A. (2006) Independent effects of socioeconomic status and place of residence on the incidence of childhood type 1 diabetes in Western Australia. *Pediatric Diabetes,* 7**,** 94-100.

HAYNES, R., DARAS, K., READING, R. & JONES, A. (2007b) Modifiable neighbourhood units, zone design and residents' perceptions. *Health & Place,* 13**,** 812-825.

HAYNES, R., READING, R. & GALE, S. (2003) Household and neighbourhood risks for injury to 5-14 year old children. *Social Science and Medicine,* 57**,** 625-636.

HEATH, C. W. (2005) Community clusters of childhood leukemia and lymphoma: Evidence of infection? *American Journal of Epidemiology,* 162**,** 817-822.

HEATH, C. W. J. (1996) Epidemiology and hereditary aspects of acute leukaemia. IN WIERNIK, P. H., CANELLOS, G. P., DUTCHER, J. P. & KYLE, R. A. (Eds.) *Neoplastic Diseases of the Blood.* New York, Library of Congress.

HEATH, C. W. J. & HASTERLIK, R. J. (1963) Leukemia among children in a suburban community. *American Journal of Medicine,* 34**,** 796-812.

HEENAN, L. D. B. (1979) Internal migration: inventory and appraisal. IN NEVILLE, R. J. W. & O'NEILL, C. J. (Eds.) *The Population of New Zealand.* Auckland, Longman Paul.

HEENAN, L. D. B. (1985) Population redistribution and internal migration. IN UNITED NATIONS, E. A. S. C. F. A. A. T. P. (Ed.) *Population of New Zealand.* New York, United Nations.

HELLER, S., KOZLOVSKI, P. & KURTZHALS, P. (In press) Insulin's 85th anniversary - an enduring medical miracle. *Diabetes Research and Clinical Practice,* Corrected Proof.

HEWITT, D. (1955) Some features of leukaemia mortality. *British Journal of Preventive Social Medicine,* 9**,** 81-88.

HIJIYA, N., HUDSON, M. M., LENSING, S., ZACHER, M., ONCIU, M., BEHM, F. G., RAZZOUK, B. I., RIBEIRO, R. C., RUBNITZ, J. E., SANDLUND, J. T., RIVERA, G. K., EVANS, W. E., RELLING, M. V. & PUI, C.-H. (2007) Cumulative Incidence of Secondary Neoplasms as a First Event After Childhood Acute Lymphoblastic Leukemia. *JAMA,* 297**,** 1207-1215.

HJALGRIM, L. L., ROSTGAARD, K., SCHMIEGELOW, K., SÖDERHÄLL, S., KOLMANNSKOG, S., VETTENRANTA, K., KRISTINSSON, J., CLAUSEN, N., MELBYE, M., HJALGRIM, H. & GUSTAFSSON, G. (2003) Age- and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. *Journal of the National Cancer Institute,* 95**,** 1539-1544.

HJALMARS, U. & GUSTAFSSON, G. (1999) Higher risk for acute childhood lymphoblastic leukaemia in Swedish population centres 1973-1994. *British Journal of Cancer,* 79**,** 30-33.

HJALMARS, U., KULLDORFF, M., GUSTAFSSON, G. & NAGARWALLA, N. (1996) Childhood leukaemia in Sweden: Using GIS and a spatial scan statistic for cluster detection. *Statistics in Medicine,* 15**,** 707-715.

HO, E. S. (2004) Mental health of Asian immigrants in New Zealand: A review of key issues. *Asian and Pacific Migration Journal,* 13**,** 39-60.

HOBBS, M., MOOR, C., WANSBROUGH, T. & CALDER, L. (2002) The health status of asylum seekers screened by Auckland Public Health in 1999 and 2000. *New Zealand Medical Journal,* 115.

HOFFBRAND, A. V., PETTIT, J. E. & MOSS, P. A. H. (2001) *Essential Haematology,* Oxford, Blackwell Science.

HONEYMAN, M. (2005) How robust is the evidence for viruses in the induction of type 1 diabetes? *Current Opinion in Immunology,* 17**,** 616-623.

HONEYMAN, M. C., COULSON, B. S., STONE, N. L., GELLERT, S. A., GOLDWATER, P. N., STEELE, C. E., COUPER, J. J., TAIT, B. D., COLMAN, P. G. & HARRISON, L. C. (2000) Association between rotavirus infection and pancreatic islet autoimmunity in children at risk of developing type 1 diabetes. *Diabetes,* 49**,** 1319.

HORWITZ, M. S., BRADLEY, L. M., HARBERTSON, J., KRAHL, T., LEE, J. & SARVENNICK, N. (1998) Diabetes induced by coxsackie virus: initiation by bystander damage and not molecular mimicry. *Nature Medicine,* 4**,** 781-785.

HORWITZ, M. S., ILIC, A., FINE, C., RODRIGUEZ, E. & SARVETNICK, N. (2002) Presented antigen from damaged pancreatic beta cells activates autoreactive T cells in virus-mediated autoimmune diabetes. *Journal Of Clinical Investigation,* 109**,** 79-87.

HUGO, G. (2004) New Zealanders in Australia in 2001. *New Zealand Population Review,* 30**,** 61-92.

HUGO, G. (2007) Population geography. *Progress in Human Geography,* 31**,** 77.

HUIZINGH (2007) *Applied statistics with SPSS,* London, Sage Publications.

HULL, D. (1979) Migration, adaptation and illness: a review. *Social Science & Medicine,* 13A**,** 25-36.

HUMMEL, M., FUCHTENBUSCH, M., SCHENKER, M. & ZIEGLER, A. G. (2000) No major association of breast-feeding, vaccinations, and childhood viral diseases with early islet autoimmunity in the German BABYDIAB study. *Diabetes Care,* 23**,** 969-974.

HUMMEL, S., WINKLER, C., SCHOEN, S., KNOPFF, A., MARIENFELD, S., BONIFACIO, E. & ZIEGLER, A. G. (2007) Breastfeeding habits in families with Type 1 diabetes. *Diabetic Medicine,* 24**,** 671-676.

HYOTY, H. & AL., E. (1988) Mumps infections in the etiology of type 1 (insulin-dependent) diabetes. *Diabetes Research,* 9**,** 111-116.

HYÖTY, H. & TAYLOR, K. (2002) The role of viruses in human diabetes. *Diabetologia,* 45**,** 1353.

HYPPONEN, E., KENWARD, M. G., VIRTANEN, S. M., PIITULAINEN, A. & ET AL. (1999) Infant feeding, early weight gain, and risk of type 1 diabetes. *Diabetes Care,* 22**,** 1961.

HYPPONEN, E., LAARA, E., REUNANEN, A., JARVELIN, M.-R. & VIRTANEN, S. M. (2001) Intake of vitamin D and risk of type 1 diabetes: a birth-cohort study. *The Lancet,* 358**,** 1500-1503.

HYPPONEN, E., VIRTANEN, S. M., KENWARD, M. G., KNIP, M. & AKERBLOM, H. K. (2000) Obesity, increased linear growth, and risk of type 1 diabetes in children. *Diabetes Care,* 23**,** 1755-1760.

HYTTINEN, V., KAPRIO, J., KINNUNEN, L., KOSKENVUO, M. & TUOMILEHTO, J. (2003) Genetic liability of type 1 diabetes and the onset age among 22, 650 young Finnish twin pairs: A nationwide follow-up study. *Diabetes,* 52**,** 1052-1055.

IMAI, S., TEZUKA, H. & FUJITA, K. (2001) A factor of inducing IgE from a filarial parasite prevents insulin-dependent diabetes mellitus in nonobese diabetic mice. *Biochemical and Biophysical Research Communications,* 286**,** 1051-1058.

INFANTE-RIVARD, C., FORTIER, I. & OLSON, E. (2000) Markers of infection, breastfeeding and childhood acute lymphoblastic leukaemia. *British Journal of Cancer,* 83**,** 1559-1564.

INFANTE-RIVARD, C., LABUDA, D., KRAJINOVIC, M. & SINNETT, D. (1999) Risk of Childhood Leukemia Associated with Exposure to Pesticides and with Gene Polymorphisms. *Epidemiology,* 10**,** 481-487.

INTERNATIONAL DIABETES FEDERATION (2005) Diabetes atlas: incidence at a glance. Brussels, International Diabetes Federation.

IZUMI, S., SUYAMA, A. & KOYAMA, K. (2003) Radiation-related mortality among offspring of atomic bomb survivors: A half-century of follow-up. *International Journal of Cancer,* 107**,** 292-297.

JARRETT, O. (1987) Leukaemogenic viruses. IN WHITTAKER, J. A. & DELAMORE, I. W. (Eds.) *Leukaemia.* Oxford, Blackwell Scientific Publications.

JATRANA, S., GRAHAM, E. & BOYLE, P. (2005) Understanding migration and health in Asia. IN JATRANA, S., TOYOTA, M. & YEOH, B. A. S. (Eds.) *Migration and health in Asia.* Abingdon, Routledge.

JOHANSSON, C., SAMUELSSON, U. & LUDVIGSSON, J. (1994) A high weight gain early in life is associated with an increased risk of type 1 (insulin-dependent) diabetes mellitus. *Diabetologia,* 37**,** 91-94.

JOHN, T. J. & SAMUEL, R. (2000) Herd immunity and herd effect: new insights and definitions. *European Journal of Epidemiology,* 16**,** 601-606.

JOHNSTON, S. L. & OPENSHAW, P. J. M. (2001) The protective effect of childhood infections. *British Medical Journal,* 322**,** 376-378.

JONES, K. & DUNCAN, C. (1995) Individuals and their ecologies: analysing the geography of chronic illness within a multilevel modelling framework. *Health & Place,* 1**,** 27-40.

JOURDAN-DA SILVA, N. & *AL., E.* (2004) Infectious diseases in the first year of life, perinatal characteristics and childhood acute leukaemia. *British Journal of Cancer,* 90**,** 139-145.

JUN, H. S. & YOON, J. W. (2001) The role of viruses in type I diabetes: two distinct cellular and molecular pathogenic mechanisms of virus-induced diabetes in animals. *Diabetologia,* 44**,** 271.

JUTO, P. (1985) Human milk stimulates beta cell function. *Archives of Disease in Childhood,* 60**,** 610–3.

KAILA, B. & TABACK, S. P. (2001) The effect of day care exposure on the risk of developing type 1 diabetes: a meta-analysis of case-control studies.
. *Diabetes Care* 24**,** 1353–58.

KALETSCH, U., KAATSCH, P., MEINERT, R., SCHÜZ, J., CZARWINSKI, R. & MICHAELIS, J. (1999) Childhood cancer and residential radon exposure – results of a population-based case-control study in Lower Saxony (Germany). *Radiation and Environmental Biophysics,* 38**,** 211-215.

KALIPENI, E. & OPPONG, J. (1998) The refugee crisis in Africa and implications for health and disease: a political ecology approach. *Social Science & Medicine,* 46**,** 1637-1653.

KAMPER-JORGENSEN, M., WOODWARD, A., WOHLFAHRT, J., BENN, C. S., SIMONSEN, J., HJALGRIM, H. & SCHMIEGELOW, K. (2008) Childcare in the first 2 years of life reduces the risk of childhood acute lymphoblastic leukemia. *Leukemia,* 22**,** 189-193.

KARGES, W., HAMMOND-MCKIBBEN, D., CHEUNG, R. K., VISCONTI, M. & ET AL. (1997) Immunological aspects of nutritional diabetes prevention in NOD mice: A pilot study for the cow's milk-based IDDM prevention trial. *Diabetes,* 46**,** 557.

KARVONEN, M., JANTTI, V., MUNTONI, S., STABILINI, M., STABILINI, L., MUNTONI, S. & TUOMILEHTO, J. (1998) Comparison of the seasonal pattern in the clinical onset of NIDDM in Finland and Sardinia. *Diabetes Care,* 21**,** 1101-1109.

KARVONEN, M., PITKANIEMI, J. & TUOMILEHTO, J. (1999) The onset age of type 1 diabetes in Finnish children has become younger. *Diabetes Care,* 22**,** 1066.

KARVONEN, M., PITKANIEMI, M., PITKANIEMI, J., KOHTAMAKI, K., TAJIMA, N., TUOMILEHTO, J., TUOMILEHTO, J., DABEE, J., KARVONEN, M., DOWSE, G. K., GAREBOO, H., VIRTALA, E. & TIIHONEN, M. (1997a) Sex difference in the incidence of insulin-dependent diabetes mellitus: An analysis of the recent epidemiological data. *Diabetes/Metabolism Reviews,* 13**,** 275-291.

KARVONEN, M., RUSANEN, J., SUNDBERG, M., VIRTALA, E., COLPAERT, A., NAUKKARINEN, A. & TUOMILEHTO, J. (1997b) Regional differences in the incidence of insulin-dependent diabetes mellitus among children in Finland from 1987 to 1991. Childhood Diabetes in Finland (DiMe) Study Group. *Annals of Medicine,* 29**,** 297-304. .

KARVONEN, M., SEKIKAWA, A., LAPORTE, R., TUOMILEHTO, J. & TUOMILEHTO-WOLF, E. (2001) Type 1 diabetes: global epidemiology. IN EKOE, J. M., ZIMMET, P. Z. & WILLIAMS, R. (Eds.) *The epidemiology of diabetes mellitus: an international perspective.* Chichester, John Wiley and Sons.

KARVONEN, M., TUOMILEHTO, J. & PODAR, T. (2003) Epidemiology of type 1 diabetes. IN PICKUP, J. C. & WILLIAMS, G. (Eds.) *Textbook of diabetes 1.* 3rd ed. Oxford, Blackwell Publishing.

KARVONEN, M., VIIK-KAJANDER, M., MOLTCHANOVA, E., LIBMAN, I., LAPORTE, R. & TUOMILEHTO, J. (2000) Incidence of childhood type 1 diabetes worldwide. *Diabetes Care,* 23**,** 1516-1526.

KEARNS, R. & MOON, G. (2002) From medical to health geography: novelty, place and theory after a decade of change. *Progress in Human Geography,* 26**,** 605-625.

KEARNS, R. A. (1993) Place and health: towards a reformed medical geography? *The Professional Geographer,* 45**,** 139-147.

KEARNS, R. A. & GESLER, W. M. (Eds.) (1998) *Putting health into place: landscape, identity and well-being,* Syracuse, Syracuse University Press.

KIBIRIGE, M., RENUKA, R., METCALF, B. & WILKIN, T. J. (2003) Testing the accelerator hypothesis: the relationship between body mass and age at diagnosis of type 1 diabetes. *Diabetes Care,* 26**,** 2865–2870.

KIMPIMÄKI, T., ERKKOLA, M., KORHONEN, S., KUPILA, A., VIRTANEN, S. M., ILONEN, J., SIMELL, O. & KNIP, M. (2001) Short-term exclusive breastfeeding predisposes young children with increased genetic risk of type I diabetes to progressive beta-cell autoimmunity. *Diabetologia,* 44**,** 63-69.

KING, L. J. (1969) *Statistical analysis in geography,* New Jersey, Prentice Hall.

KING, R. (1995) Migrations, globalisation and place. IN MASSEY, D. S. & JESS, P. (Eds.) *A Place in the World?* Oxford, The Open University.

KINGHAM, S., FISHER, G., HALES, S., WILSON, I. & BARTIE, P. (2008) An empirical model for estimating census unit population exposure in areas lacking air quality monitoring. *Journal of Exposure Science and Environmental Epidemiology,* 18**,** 200-210.

KINLEN, L. (2004a) Infections and immune factors in cancer: the role of epidemiology. *Oncogene,* 23**,** 6341-6348.

KINLEN, L. J. (1988) Evidence for an infective cause of childhood leukaemia: a comparison of a Scottish new town with nuclear reprocessing sites in Britain. *The Lancet* 1323-1327.

KINLEN, L. J. (1997) High-contact paternal occupations, infection and childhood leukaemia: five studies of unsual population mixing of adults. *British Journal of Cancer,* 76**,** 1539-1545.

KINLEN, L. J. (2000) Infection, childhood leukaemia and the Seascale cluster. *Radiological Protection Bulletin,* 226**,** 9-18.

KINLEN, L. J. (2004b) RE: Childhood cancer and population mixing. *American Journal of Epidemiology,* 159**,** 716.

KINLEN, L. J. (2006) Childhood leukaemia and ordnance factories in west Cumbria during the Second World War. *British Journal of Cancer,* 95 102-106.

KINLEN, L. J. & BALKWILL, A. (2001) Infective cause of childhood leukaemia and wartime population mixing in Orkney and Shetland,UK. *The Lancet,* 357**,** 858.

KINLEN, L. J. & BRAMALD, S. (2001) Paternal occupational contact level and childhood leukaemia in rural Scotland: a case-control study. *British Journal of Cancer,* 84**,** 1002-1007.

KINLEN, L. J., CLARKE, K. & HUDSON, C. (1990) Evidence from population mixing in British new towns 1946-1985 of an infective basis for childhood leukaemia. *The Lancet,* 336**,** 577-82.

KINLEN, L. J., DICKSON, M. & STILLER, C. A. (1995) Childhood leukaemia and non-Hodgkin's lymphoma near large rural construction sites, with a comparison with Sellafield nuclear site. *British Medical Journal,* 310**,** 763-768.

KINLEN, L. J. & DOLL, R. (2004) Population mixing and childhood leukaemia: Fallon and other US clusters. *British Journal of Cancer,* 91**,** 1-3.

KINLEN, L. J. & HUDSON, C. (1991) Childhood leukaemia and poliomyelitis in relation to military encampments in England and Wales in the period of national military service, 1950-63. *British Medical Journal,* 303**,** 1357-1362.

KINLEN, L. J., HUDSON, C. M. & STILLER, C. A. (1991) Contacts between adults as evidence for an infective origin in childhood leukaemia: an explanation for the excess near nuclear establishments in West Berkshire? *British Journal of Cancer,* 64**,** 549-554.

KINLEN, L. J., JIANG, J. & HEMMINKI, K. (2002) A case-control study of childhood leukaemia and paternal occupational contact level in rural Sweden. *British Journal of Cancer,* 86**,** 732-737.

KINLEN, L. J. & JOHN, S. M. (1994) Wartime evacuation and mortality from childhood leukaemia in England and Wales in 1945-9. *British Medical Journal,* 309**,** 1197-1201.

KINLEN, L. J., O'BRIEN, F., CLARKE, K., BALKWILL, A. & MATTHEWS, F. (1993) Rural population mixing and childhood leukaemia: effects of the North Sea oil industry in Scotland, including the area near Dounreay nuclear site. *British Medical Journal,* 306**,** 743-8.

KINLEN, L. J. & PETRIDOU, E. (1995) Childhood leukaemia and rural population movements: Greece, Italy and other countries. *Cancer Causes and Control,* 6**,** 445-50.

KNOX, E. (1989) Detection of clusters. IN ELLIOTT, P. (Ed.) *Methodology of enquiries into disease clustering* London, Small Area Health Statistics Unit.

KOLB, H. & ELLIOT, R. B. (1994) Increasing incidence of IDDM a consequence of improved hygiene? *Diabetologia,* 37**,** 729-731.

KOSTRABA, J. N., GAY, E. C., REWERS, M. & HAMMAN, R. F. (1992) Nitrate levels in community drinking waters and risk of IDDM. An ecological analysis. *Diabetes Care,* 15**,** 1505-1508.

KOUSHIK, A., KING, W. D. & MCLAUGHLIN, J. R. (2001) An ecologic study of childhood leukemia and population mixing in Ontario, Canada. *Cancer Causes and Control,* 12**,** 483-490.

KRAVETS, N. & HADDEN, W. C. (2007) The accuracy of address coding and the effects of coding errors. *Health and Place,* 13**,** 293-298

KREMER, H. U. (1947) Juvenile diabetes as a sequel to mumps. *American Journal of Medicine,* 3**,** 257–259.

KUEHNI, C. E. & ZWAHLEN, M. (2006) Commentary: Numerous, heterogeneous, and often poor - the studies on childhood leukaemia and socioeconomic status. *International Journal of Epidemiology,* 35**,** 384-385.

KULLDORFF, M. & INFORMATION MANAGEMENT SERVICES, I. (2002) *SaTScan v.3.05: Software for the spatial and space-time scan statistic,* Bethesda, National Cancer Institute.

KULLDORFF, M. A. (1997) A spatial scan statistic. *Communications in Statistics: Theory and Methods,* 26**,** 1481-1496.

KULLDORFF, M. A. (2006) SaTScan user guide for version 6.1. Rockville, MD Information Management Services Inc.

KUMAR, N. (2004) Changing geographic access to and locational efficiency of health services in two Indian districts between 1981 and 1996. *Social Science and Medicine,* 58**,** 2045-2067.

KWAN, M. L., BUFFLER, P. A., ABRAMS, B. & KILEY, V. A. (2004) Breastfeeding and the risk of childhood leukemia: a meta-analysis. *Public Health Reports,* 119**,** 521-535.

LABAR, B., RUDAN, I., IVANKOVIC, D., BILOGLAV, Z., MRSIC, M., STRNAD, M., FUCIC, A., ZNAOR, A., BRADIC, T. & CAMPBELL, H. (2004) Haematological malignancies in childhood in Croatia: Investigating the theories of depleted uranium, chemical plant damage and 'population mixing'. *European Journal of Epidemiology,* 19**,** 55-60.

LAMB, W. H. (2006) Diabetes mellitus, type 1. eMedicine WebMD.

LAMMI, N., KARVONEN, M. & TUOMILEHTO, J. (2005) Do microbes have a causal role in type 1 diabetes? *Medical Science Monitor,* 11**,** RA63-69.

LANGFORD, I. (1991) Childhood leukaemia mortality and population change in England and Wales 1969-73. *Social Science & Medicine,* 33**,** 435-440.

LANGFORD, I. H. & BENTHAM, G. (1993) Epidemiology of childhood leukemia. *British Medical Journal,* 307**,** 445–446.

LANGHOLZ, B., EBI, K., THOMAS, C., PETERS, J. & LONDON, S. (2002) Traffic density and the risk of childhood leukemia in a Los Angeles case-control study. *Annals of Epidemiology,* 12**,** 482–487.

LARON, Z. (2002) Incidence and seasonality of type 1 diabetes mellitus - What now? *Journal of Pediatric Endocrinology and Metabolism,* 15**,** 573-575.

LARON, Z., LEWY, H., WILDERMAN, I., CASU, A., WILLIS, J., REDONDO, M. J., LIBMAN, I., WHITE, N. & CRAIG, M. (2005) Seasonality of month of birth of children and adolescents with type 1 diabetes mellitus in homogenous and heterogeneous populations. *Israel Medical Association Journal,* 7**,** 381-384.

LAW, G. R. (2005) Invited commentary: Do clusters of leukemia and lymphoma provide evidence for an infectious cause?' *American Journal of Epidemiology,* 162**,** 823-824.

LAW, G. R., MCKINNEY, P. A., STAINES, A., WILLIAMS, R. & ET AL. (1997) Clustering of childhood IDDM: Links with age and place of residence. *Diabetes Care,* 20**,** 753.

LAW, G. R., PARSLOW, R. C. & ROMAN, E. (2003) Childhood cancer and population mixing. *American Journal of Epidemiology,* 158**,** 328–336.

LAWLER-HEAVNER, J., CRUICKSHANKS, K. J., HAMMAN, R. F., GAY, E. C., KLINGENSMITH, G. & CHASE, H. P. (1991) Household density in early childhood and risk of insulin-dependent diabetes (IDDM). *Diabetes,* 40**,** 319A.

LAWSON, A. B. (2001) *Statistical methods in spatial epidemiology* Chichester, John Wiley.

LEHTINEN, M., KOSKELA, P., ÖGMUNDSDOTTIR, H. M., BLOIGU, A., DILLNER, J., GUDNADOTTIR, M., HAKULINEN, T., KJARTANSDOTTIR, A., KVARNUNG, M., PUKKALA, E., TULINIUS, H. & LEHTINEN, T. (2003) Maternal herpesvirus infections and risk of acute lymphoblastic leukemia in the offspring. *American Journal of Epidemiology,* 158**,** 207-213.

LESLIE, R. D. & ELLIOTT, R. B. (1994) Early environmental events as a cause of IDDM. Evidence and implications. *Diabetes,* 43**,** 843-850.

LEUKAEMIA FOUNDATION (2003) *Understanding leukaemia and related bone marrow disorders,* Mosman, Leukaemia Foundation.

LÉVY-MARCHAL, C., PATTERSON, C. & GREEN, A. (1995) Variation by age group and seasonality at diagnosis of childhood IDDM in Europe. The EURODIAB ACE Study Group. *Diabetologia,* 38**,** 823-30.

LI, C. Y., LIN, R. S. & LIN, C. H. (1998) Urbanization and childhood leukaemia in Taiwan. *International Journal of Epidemiology,* 27**,** 587-591.

LIANG, D.-C. & PUI, C.-H. (2005) Childhood acute lymphoblastic leukaemia. IN HOFFBRAND, A. V., CATOVSKY, D. & TUDDENHAM, E. G. D. (Eds.) *Postgraduate Haematology.* 5th edition ed. Massachusetts, Blackwell Publishing.

LIESNER, R. J. & GOLDSTONE, A. H. (1997) ABC of clinical haematology: the acute leukaemias. *British Medical Journal,* 314**,** 733-740.

LIKE, A. A., GUBERSKI, D. L. & BUTLER, L. (1991) Influence of environmental viral agents on frequency and tempo of diabetes mellitus in BB/Wor rats. *Diabetes,* 40**,** 259-62.

LITTLE, J. (1999) *Epidemiology of childhood cancer: IARC scientific publications no. 149,* Lyon, International Agency for Research on Cancer.

LITVA, A. & EYLES, J. (1995) Coming out: exposing social theory in medical geography. *Health & Place,* 1**,** 5-14.

LONG, J. S. (1997) *Regression models for categorical and limited dependent variables* London Sage Publications.

LONGHURST, R. (1994) The geography closest in - the body... the politics of pregnability. *Australian Geographical Studies,* 32**,** 214-223.

LONGINI, I. M. J., FINE, P. E. M. & THACKER, S. B. (1986) Predicting the global spread of new infectious agents. *American Journal of Epidemiology,* 123**,** 383-391.

LOVETT, A., BENTHAM, G. & FLOWERDEW, R. (1986) Analysing geographic variations in mortality using poisson regression: the example of ischaemic heart diseases in England and Wales 1969-1973. *Social Science & Medicine,* 23**,** 935-943.

LOVETT, A. & FLOWERDEW, R. (1989) Analysis of count data using poisson regression. *Professional Geographer,* 41**,** 190-198.

LUBIN, J. H., LINET, M. S., BOICE, J. D., BUCKLEY, J. & ET AL. (1998) Case-control study of childhood acute lymphoblastic leukemia and residential radon exposure. *Journal of the National Cancer Institute,* 90**,** 294.

MA, X., BUFFLER, P. A., GUNIER, R. B., DAHL, G., SMITH, M. T., REINIER, K. & REYNOLDS, P. (2002a) Critical windows of exposure to household pesticides and risk of childhood leukemia. *Environmental Health Perspectives,* 110**,** 955-960.

MA, X., BUFFLER, P. A., SELVIN, S., MATTHAY, K. K., WIENCKE, J. K., WIEMELS, J. L. & REYNOLDS, P. (2002b) Daycare attendance and risk of childhood acute lymphoblastic leukaemia. *British Journal of Cancer,* 86**,** 1419-1424.

MA, X., BUFFLER, P. A., WIEMELS, J. L., SELVIN, S., METAYER, C., LOH, M., DOES, M. B. & WIENCKE, J. K. (2005) Ethnic difference in daycare attendance, early infections, and risk of childhood acute lymphoblastic leukemia. *Cancer Epidemiology Biomarkers & Prevention,* 14**,** 1928-1934.

MACFARLANE, A. J. & SCOTT, F. W. (2003) Environmental agents and type 1 diabetes. IN PICKUP, J. C. & WILLIAMS, G. (Eds.) *Textbook of diabetes 1*. Oxford, Blackwell Publishing.

MACMAHON, B. & TRICHOPOULOS, D. (1996) *Epidemiology: Principles & Methods,* Boston, Little, Brown & Company.

MAGURRAN, A. E. (1988) *Ecological diversity and its measurement* London, Croom Helm.

MANLEY, D., FLOWERDEW, R. & STEEL, D. (2006) Scales, levels and processes: studying spatial patterns of British census variables. *Computers, Environment and Urban Systems,* 30**,** 143-160.

MASSEY, D. S., ARANGO, J., HUGO, G., KOUAOUCI, A., PELLEGRINO, A. & TAYLOR, J. E. (1998) *Worlds in motion: understanding international migration at the end of the millenium,* Oxford, Clarendon Press.

MASSEY, D. S., ARANGO, J., HUGO, G. & TAYLOR, J. E. (1993) Theories of international migration: a review and apprasial. *Population and Development Review,* 19**,** 431-466.

MATHIEU, C. & BADENHOOP, K. (2005) Vitamin D and type 1 diabetes mellitus: state of the art. *Trends in Endocrinology and Metabolism,* 16**,** 261-266.

MATHIEU, C., WAER, M., LAUREYS, J., RUTGEERTS, O. & BOUILLON, R. (1994) Prevention of autoimmune diabetes in NOD mice by 1,25 dihydroxyvitamin D3. *Diabetologia,* 37**,** 552-558.

MATTHEWS, M. (1993) Radioactive fallout in the South Pacific: a History. Part 3: strontium-90 and caesium-137 deposition in New Zealand and resulting contamination of milk. *National Radiation Laboratory Report.* Wellington, Minisitry of Health.

MAUGH, T. H. (1975) Diabetes: epidemiology suggests a viral connection. *Science,* 188**,** 347-351.

MAULE, M. M., ZUCCOLO, L., MAGNANI, C., PASTORE, G., DALMASSO, P., PEARCE, N., MERLETTI, F. & GREGORI, D. (2006) Bayesian methods for early detection of changes in childhood cancer incidence: Trends for acute lymphoblastic leukaemia are consistent with an infectious aetiology. *European Journal of Cancer,* 42**,** 78-83.

MAYER, J. D. & MEADE, M. S. (1994) A reformed medical geography reconsidered. *The Professional Geographer,* 46**,** 103-106.

MCBRIDE, M. L., GALLAGHER, R. P., THERIAULT, G., ARMSTRONG, B. G., TAMARO, S., SPINELLI, J. J., DEADMAN, J. E., FINCHAM, S., ROBSON, D. & CHOI, W.

(1999) Power-frequency electric and magnetic fields and risk of childhood leukemia in Canada. *American Journal of Epidemiology,* 149**,** 831-842.

MCCREDIE, M., WILLIAMS, S. & COATES, M. (1999) Cancer mortality in East and Southeast Asian migrants to New South Wales, Australia, 1975-1995. *British Journal of Cancer,* 79**,** 1277-1282.

MCHUGH, K. E. (2000) Inside, outside, upside down, backward, forward, round and round: a case for ethnographic studies in migration. *Progress in Human Geography,* 24**,** 71-89.

MCKAY, L., MACINTYRE, S. & ELLAWAY, A. (2003) Migration and health: a review of the international literature. IN UNIT, M. R. C. S. A. P. H. S. (Ed.) *Occasional Paper No 12.* Glasgow, University of Glasgow.

MCKINNEY, P. A. & EURODIAB SEASONALITY OF BIRTH GROUP (2001) Seasonality of birth in patients with childhood type 1 diabetes in 19 European regions. *Diabetologia,* 44**,** B67-B74.

MCKINNEY, P. A., JUSZCZAK, E., FINDLAY, E., SMITH, K. & THOMSON, C. S. (1999a) Pre- and perinatal risk factors for childhood leukaemia and other malignancies: a Scottish case control study. *British Journal of Cancer,* 80**,** 1844-1851.

MCKINNEY, P. A., LAW, G. R., BODANSKY, H. J., STAINES, A. & WILLIAMS, D. R. R. (1996) Geographical mapping of childhood diabetes in the northern English county of Yorkshire. *Diabetic Medicine,* 13**,** 734-40.

MCKINNEY, P. A., OKASHA, M., PARSLOW, R. C., LAW, G. R., GURNEY, K. A. W., WILLIAMS, R. & BODANSKY, H. J. (2000) Early social mixing and childhood Type 1 diabetes mellitus: a case-control study in Yorkshire, UK. *Diabetic Medicine,* 17**,** 236-242.

MCKINNEY, P. A., PARSLOW, R., GURNEY, K. A., LAW, G. R. & ET AL. (1999b) Perinatal and neonatal determinants of childhood type 1 diabetes: A case-control study in Yorkshire, U.K. *Diabetes Care,* 22**,** 928.

MCNALLY, R., FELTBOWER, R., PARKER, L., BODANSKY, H., CAMPBELL, F. & MCKINNEY, P. (2006a) Space–time clustering analyses of type 1 diabetes among 0- to 29-year-olds in Yorkshire, UK. *Diabetologia,* 49**,** 900-904.

MCNALLY, R. J. Q., ALEXANDER, F. E. & BIRCH, J. M. (2002) Space-time clustering analyses of childhood acute lymphoblastic leukaemia by immunophenotype. *British Journal of Cancer,* 87**,** 513-515.

MCNALLY, R. J. Q., ALEXANDER, F. E. & BITHELL, J. F. (2006b) Space-time clustering of childhood cancer in Great Britain: A national study, 1969-1993. *International Journal of Cancer,* 118**,** 2840-2846.

MCNALLY, R. J. Q., ALSTON, R. D., CAIRNS, D. P., EDEN, T. O. B. & BIRCH, J. M. (2003) Geographical and ecological analyses of childhood acute leukaemias and lymphomas in north-west England. *British Journal of Haematology,* 123**,** 60-65.

MCNALLY, R. J. Q. & EDEN, T. O. B. (2004) An infectious aetiology for childhood acute leukaemia: a review of the evidence. *British Journal of Haematology,* 127**,** 243-263.

MCNALLY, R. J. Q. & PARKER, L. (2006) Environmental factors and childhood acute leukemias and lymphomas *Leukemia & Lymphoma,* 47 583-598

MCNAMEE, R. (2003) Confounding and confounders. *Occupational and Environmental Medicine,* 60**,** 227-234.

MCWHIRTER, W. R. (1982) The relationship of incidence of childhood lymphoblastic leukaemia to social class. *British Journal of Cancer,* 46**,** 640-645.

MCWHIRTER, W. R. & BACON, J. E. (1980) Epidemiology of acute lymphoblastic leukaemia of childhood in Brisbane. *Medical Journal of Australia,* 2**,** 154-5.

METGE, J. (2004) *A new Māori migration: rural and urban relations in northern New Zealand,* Oxford, Berg.

MILLER, J. E. (2005) *The Chicago guide to writing about multivariate analysis* Chicago, University of Chicago Press.

MILLER, L. J., PEARCE, J. & BARNETT, R. (2007) The place of population mixing in the aetiology of disease: a New Zealand perspective. *Australasian Epidemiologist,* 14**,** 12–15.

MILLER, L. J., PEARCE, J., BARNETT, R., WILLIS, J. A., DARLOW, D. A. & SCOTT, R. S. (in review) Is population mixing associated with childhood type 1 diabetes in Canterbury, New Zealand? *Social Science & Medicine*.

MILNE, E., LAURVICK, C. L., DE KLERK, N., ROBERTSON, L., THOMPSON, J. R. & BOWER, C. (2007) Trends in childhood acute lymphoblastic leukemia in Western Australia, 1960-2006. *International Journal of Cancer,* In press.

MINISTRY OF EDUCATION (2007) Schooling in New Zealand - a guide. IN MINISTRY OF EDUCATION (Ed.).

MINISTRY OF HEALTH (2002) Breastfeeding: a guide to action. Wellington, Ministry of Health.

MINISTRY OF HEALTH (2004) Tracking the Obesity Epidemic: New Zealand 1977–2003 Wellington., Ministry of Health.

MINISTRY OF HEALTH (2005) Monitoring health inequality through neighbourhood life expectancy. Public Health Intelligence occasional bulletin no. 28. Wellington, Ministry of Health.

MINISTRY OF HEALTH, PARNELL W, SCRAGG R & AL., E. (2003) Key results of the 2002 National Children's Nutrition Survey. Wellington: Ministry of Health. Wellington, Ministry of Health.

MINISTRY OF HEALTH, M. (1998) Our children's health: key findings on the health of New Zealand children. Wellington, Ministry of Health.

MINISTRY OF SOCIAL DEVELOPMENT (2002) New Zealand's agenda for children - making life better for children. Wellington, Ministry of Social Development.

MINISTRY OF TOURISM (2007) International visitor survey. Wellington, Ministry of Tourism.

MIRDAL, G. M. (1984) Stress and distress in migration - problems and resources of Turkish women in Denmark. *International Migration Review,* 18**,** 984-1003.

MIRDAL, G. M. (2006) Stress and distress in migration: twenty years after. *International Migration Review,* 40**,** 375-389.

MIRSKY, J., SLONIM-NEVO, V. & RUBINSTEIN, L. (2007) Psychological wellness and distress among recent immigrants: A four-year longitudinal study in Israel and Germany. *International Migration,* 45**,** 151-175.

MOON, G. & BARNETT, R. (2003) Spatial scale and the geography of tobacco smoking in New Zealand: a multilevel perspective. *New Zealand Geographer,* 59**,** 6-15.

MOON, G., GOULD, M. & *AL., E.* (2000) *Epidemiology: An introduction,* Buckingham, Open University Press.

MORADI, T., DELFINO, R. J., BERGSTRÖM, S. R., YU, E. S., ADAMI, H. O. & YUEN, J. (1998) Cancer risk among Scandinavian immigrants in the US and Scandinavian residents compared with US whites, 1973-89. *European Journal of Cancer Prevention,* 7**,** 117-25.

MORI, H., COLMAN, S. M., XIAO, Z., FORD, A. M., HEALY, L. E., DONALDSON, C., HOWS, J. M., NAVARRETE, C. & GREAVES, M. (2002) Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proceedings of the National Academy of Sciences,* 99**,** 8242-8247.

MUIRHEAD, C. R. (1995) Childhood leukemia in metropolitan regions in the United States: a possible relation to population density? *Cancer Causes and Control,* 6**,** 383-388.

MURRAY, L., MCCARRON, P., BAILIE, K., MIDDLETON, R., DAVEY-SMITH, G., DEMPSEY, S., MCCARTHY, A. & GAVIN, A. (2002) Association of early life factors and acute lymphoblastic leukaemia in childhood: historical cohort study. *British Journal of Cancer,* 86**,** 356-361.

NASTERLACK, M. (2006) Do pesticides cause childhood cancer? *International Archives of Occupational and Environmental Health,* 79 536-544.

NASTERLACK, M. (2007) Pesticides and childhood cancer: An update. *International Journal of Hygiene and Environmental Health,* 210**,** 645-657.

NATIONAL RADIATION LABORATORY (1998) Sources, effects and risks of ionising radiation. *Information Sheet 5.* Wellington, National Radiation Laboratory.

NAUMBURG, E., ., BELLOCCO, R., CNATTINGIUS, S., JONZON, A. & EKBOM, A. (2002) Perinatal exposure to infection and risk of childhood leukemia. *Medical and Pediatric Oncology,* 38**,** 391-397.

NEGLIA, J. P., LINET, M. S., SHU, X. O., SEVERSON, R. K., POTTER, J. D., MERTENS, A. C., WEN, W., KERSEY, J. H. & ROBISON, L. L. (2000) Patterns of infection and day care utilization and risk of childhood acute lymphoblastic leukaemia. *British Journal of Cancer,* 82**,** 234-240.

NEWELL, J. (2002) Internal and international migration in New Zealand's regions, 1986-1996. IN CARMICHAEL, G. A. & DHARMALINGAM, A. (Eds.) *Populations of New Zealand and Australia at the Millenium: A joint special issue of the Journal of Population Research and the New Zealand Population Review*. Canberra and Wellington, Australian Population Association and Population Association of New Zealand.

NEWMAN, E. & SELM, J. (2003) *Refugees and forced displacement: international security, human vulnerability and the state,* New York, United Nations University Press.

NORMAN, P., BOYLE, P. & REES, P. (2005) Selective migration, health and deprivation: a longitudinal analysis. *Social Science & Medicine,* 60**,** 2755–2771.

NORRIS, J. M. & SCOTT, F. W. (1996) A meta-analysis of infant diet and insulin-dependent diabetes mellitus: do biases play a role? . *Epidemiology,* 7**,** 87-92.

NYARI, T. A., KAJTA, R., BARTYIK, K., THURZO, L. & PARKER, L. (2006) Childhood acute lymphoblastic leukaemia in relation to population mixing around the time of birth in south hungary. *Pediatric Blood & Cancer,* 47**,** 944–948.

NYSTAD, W., SKRONDAL, A. & MAGNUS, P. (1999) Day care attendance, recurrent respiratory tract infections and asthma. *Int. J. Epidemiol.,* 28**,** 882-887.

NYSTROM, L., DAHLQUIST, G., OSTMAN, J. & *AL., E.* (1992) Risk of developing insulin-dependent diabetes mellitus (IDDM) before 35 years of age: indications of climatological determinants for age at onset. *International Journal of Epidemiology,* 21**,** 352-358.

NZHIS (2004) New Zealand Cancer Registry: data dictionary version 1.2. Wellington, Ministry of Health.

O'BRIEN, M., JONES, D., SLOAN, D. & RUSTIN, M. (2000) Children's independent spatial mobility in the urban public realm. *Childhood,* 7**,** 257-277.

O CONNELL, M. A., DONATH, S. & CAMERON, F. J. (2007) Major increase in type 1 diabetes; no support for the accelerator hypothesis. *Diabetic Medicine,* 24**,** 920-923.

OHSUGI, T. & KUROSAWA, T. (1994) Increased incidence of diabetes mellitus in specific pathogen-eliminated offspring produced by embryo transfer in NOD mice with low incidence of the disease. *Laboratory Animal Science,* 44**,** 386–388.

OLDSTONE, M. B. A. (Ed.) (2005) *Molecular mimicry, microbial infection, and autoimmune disease: evolution of the concept,* Berlin, Springer-Verlag.

OLSEN, S. F., MARTUZZI, M. & ELLIOTT, P. (1996) Cluster analysis and disease mapping - why, when and how? A step by step guide. *British Medical Journal,* 313**,** 863-866.

ONGLEY, P. & PEARSON, D. (1995) Post-1945 international migration: New Zealand, Australia and Canada compared. *International Migration Review,,* 29**,** 765-793.

OPENSHAW, S. (1984) *The modifiable areal unit problem. Concepts and techniques in modern geography 38,* Norwich, Geo Books.

ØSTBYE, T., WELBY, T. J., PRIOR, I. A. M., SALMOND, C. E. & STOKES, Y. M. (1989) Type 2 (non-insulin-dependent) diabetes mellitus, migration and westernisation: The Tokelau Island migrant study. *Diabetologia,* 32**,** 585-590.

OSTERHOLM, M. T. (1994) Infectious disease in child day care: An overview. *Pediatrics,* 94**,** 987.

PARKER, L., CRAFT, A. W., SMITH, J., DICKINSON, H. O., WAKEFORD, R., BINKS, K., MCELVENNY, D., SCOTT, L. & SLOVAK, A. (1993) Geographical distribution of preconceptional radiation doses to fathers employed at the Sellafield nuclear installation, West Cumbria. . *British Medical Journal,* 307**,** 966–971.

PARKIN, D., WHELAN, S., FERLAY, J., TEPPO, L. & THOMAS, D. (Eds.) (2003) *Cancer incidence in five continents, vol. VIII* Lyon, International Agency for Research on Cancer.

PARKIN, D. M., KRAMAROVA, E., DRAPER, G. J., MASUYER, E., MICHAELIS, J., NEGLIA, J., QURESHI, S. & STILLER, C. A. (1998) International incidence of childhood cancer: volume II. *IARC Scientific Publications 144.* New York, International Agency for Research on Cancer.

PARR, H. (2002) Medical geography: diagnosing the body in medical and health geography, 1999-2000. *Progress in Human Geography,* 26**,** 240.

PARSLOW, R. C., LAW, G. R., FELTBOWER, R., KINSEY, S. E. & MCKINNEY, P. A. (2002) Population mixing, childhood leukaemia, CNS tumours and other childhood cancers in Yorkshire. *European Journal of Cancer,* 38**,** 2033–2040.

PARSLOW, R. C., LAW, G. R., FELTBOWER, R. G. & MCKINNEY, P. A. (2005) Childhood leukaemia incidence and the population mixing hypothesis in US SEER data. *British Journal of Cancer,* 92**,** 978.

PARSLOW, R. C., MCKINNEY, P. A., LAW, G. R. & BODANSKY, H. J. (2001) Population mixing and childhood diabetes. *International Journal of Epidemiology,* 30**,** 533-538.

PARSLOW, R. C., MCKINNEY, P. A., LAW, G. R., STAINES, A., WILLIAMS, R. & BODANSKY, H. J. (1997) Incidence of childhood diabetes mellitus in Yorkshire, northern England, is associated with nitrate in drinking water: an ecological analysis. *Diabetologia,* 40**,** 550-556.

PATTERSON, C. C., CARSON, D. J. & HADDEN, D. R. (1996) Epidemiology of childhood IDDM in Northern Ireland 1989-1994: low incidence in areas with highest population density and most household crowding. *Diabetologia,* 39**,** 1063-9.

PATTERSON, C. C., CARSON, D. J., HADDEN, D. R., WAUGH, N. R. & COLE, S. K. (1994) A case-control investigation of perinatal risk factors for childhood IDDM in Northern Ireland and Scotland. *Diabetes Care,* 17**,** 376-381.

PATTERSON, C. C., DAHLQUIST, G., SOLTESZ, G., GREEN, A. & EURODIAB ACE STUDY GROUP (2001) Is childhood-onset type 1 diabetes a wealth-related disease? An ecological analysis of European incidence rates. *Diabetologia,* 44**,** B9-B16.

PATTERSON, C. C. & WAUGH, N. R. (1992) Urban/rural and deprivational differences in incidence and clustering of childhood diabetes in Scotland. *International Journal of Epidemiology,* 21**,** 108-117.

PEARCE, J., BARNETT, R. & KINGHAM, S. (2006a) Slip! Slap! Slop! cutaneous malignant melanoma incidence and social status in New Zealand, 1995-2000. *Health and Place*.

PEARCE, J. & BOYLE, P. (2005) Is the urban excess in lung cancer in Scotland explained by patterns of smoking? *Social Science & Medicine,* 60**,** 2833-2843.

PEARCE, J. & DORLING, D. (2006) The place of population change in explaining geographical inequalities in health in New Zealand. *International Journal of Epidemiology,* 35**,** 1099-1105.

PEARCE, J., HISCOCK, R., BLAKELY, T. & WITTEN, K. (2008) The contextual effects of neighbourhood access to supermarkets and convenience stores on individual fruit and vegetable consumption. *Journal of Epidemiology and Community Health,* 62**,** 198-201.

PEARCE, J., WITTEN, K. & BARTIE, P. (2006b) Neighbourhoods and health: a GIS approach to measuring community resource accessibility. *Journal of Epidemiology and Community Health,* 60**,** 389-395.

PEARCE, M. S., COTTERILL, S. J. & PARKER, L. (2004) Fathers' occupational contacts and risk of childhood leukemia and non-Hodgkin lymphoma. *Epidemiology,* 15**,** 352-356.

PEARCE, M. S., HAMMAL, D. M., DORAK, M. T., MCNALLY, R. J. Q. & PARKER, L. (2006c) Paternal occupational exposure to pesticides or herbicides as risk factors for cancer in children and young adults: a case-control study from the north of England. *Archives of Environmental and Occupational Health,* 61**,** 138 - 144.

PENROSE, M. (1970) Cat leukaemia. *British Medical Journal,* i**,** 755.

PERNICE, R., TRLIN, A., HENDERSON, A. & NORTH, N. (2000) Employment and mental health of three groups of immigrants to New Zealand. *New Zealand Journal of Psychology,* 29**,** 24-29.

PERRILLAT, F., CLAVEL, J., BARUCHEL, A., LEVERER, G., NELKEN, B., PHILIPPE, N., SCHAISON, G., SOMMELET, D., VILMER, E. & HEMON, D. (2002) Day-care, early common infections and childhood acute leukaemia: a multicentre French case-control study. *British Journal of Cancer,* 86**,** 1064-1069.

PETRIDOU, E., DALAMAGA, M., MENTIS, A., SKALKIDOU, A., MOUSTAKI, M., KARPATHIOS, T. & TRICHOPOULOS, D. (2001) Evidence on the infectious etiology of childhood leukemia: the role of low herd immunity (Greece).(Author abstract). *Cancer Causes and Control,* 12**,** 645(8).

PETRIDOU, E., KASSIMOS, D., KALMANTI, M., KOSMIDIS, H., HAIDAS, S., FLYTZANI, V., TONG, D. & TRICHOPOULOS, D. (1993) Age of exposure to infections and risk of childhood leukaemia. *British Medical Journal,* 307**,** 774.

PETROVSKY, N. & SCHATZ, D. A. (2003) The immunology of human type 1 diabetes. IN PICKUP, J. C. & WILLIAMS, G. (Eds.) *Textbook of Diabetes 1.* 3rd ed. Oxford, Blackwell Publishing.

PIERCE, M. (1936) Childhood leukemia. *Journal of Pediatrics,* 8**,** 66-95.

POBEL, D. & VIEL, J. F. (1997) Case-control study of leukaemia among young people near La Hague nuclear reprocessing plant: the environmental hypothesis revisited. *British Medical Journal,* 314**,** 101-.

POOL, D. I. (1973) The effects of the 1918 pandemic of influenza on the Māori population of New Zealand. *Bulletin of the History of Medicine* 47**,** 273-281.

POOLE, C., GREENLAND, S., LUETTERS, C., KELSEY, J. L. & MEZEI, G. (2006) Socioeconomic status and childhood leukaemia: a review. *Int. J. Epidemiol.,* 35**,** 370-384.

POOLEY, C. G. & WHYTE, I. D. (1991) Introduction: approaches to the study of migration and social change. IN POOLEY, C. G. & WHYTE, I. D. (Eds.) *Migrants, emigrants and immigrants: a social history of migration.* London, Routledge.

POPULATION AND ENVIRONMENTAL HEALTH GROUP (2006) Notifiable and other diseases In New Zealand: annual report 2006. Wellington, Institute of Environmental Science and Research Limited.

PRESTON, D. L., KUSUMI, S., TOMONAGA, M., IZUMI, S., RON, E., KURAMOTO, A., KAMADA, N., DOHY, H., MATSUI, T., NONAKA, H., THOMPSON, D. E., SODA, M. & MABUCHI, K. (1994) Cancer Incidence in Atomic Bomb Survivors. Part III: Leukemia, Lymphoma and Multiple Myeloma, 1950-1987. *Radiation Research,* 137**,** S68-S97.

PRINGLE, D. G. (2003) Cancer mortality and morbidity in County Louth. County Louth, Cooley Environment and Health Group.

PRIOR, I. (1992) Tokelau: migration and health in a small Polynesian society - a longitudinal study. IN ROBERTS, D. F., FUJIKI, N. & TORIZUKA, K. (Eds.) *Isolation, migration and health.* Cambridge, Cambridge University Press.

PRIOR, I. A. M. (1975) Migration and Health. IN STANHOPE, J. M. (Ed.) *Migration and health in New Zealand and the Pacific.* Wellington, Department of Health.

PROJECTX TECHNOLOGY LTD. (2006) ZoomIn. Accessed March 21 2007, from: http://www.zoomin.co.nz/. ProjectX Technology Ltd.

PUI, C.-H., RELLING, M. V. & DOWNING, J. R. (2004) MECHANISMS OF DISEASE: Acute Lymphoblastic Leukemia. *The New England Journal of Medicine,* 350**,** 1535.

PUNDZIUTE-LYCKA, A., DAHLQUIST, G., NYSTROM, L., ARNQVIST, H., BJORK, E., BLOHME, G. & ET AL. (2002) The incidence of Type I diabetes has not increased but shifted to a younger age at diagnosis in the 0–34 years group in Sweden 1983–1998. *Diabetologia,* 45**,** 783–791.

PUNDZIUTE-LYCKA, A., URBONAITE, B. & DAHLQUIST, G. (2000) Infections and risk of type I (insulin-dependent) diabetes mellitus in Lithuanian children. *Diabetologia,* 43**,** 1229-1234.

RAASCHOU-NIELSEN, O., HERTEL, O., THOMSEN, B. L. & OLSEN, J. H. (2001) Air pollution from traffic at the residence of children with cancer. *American Journal of Epidemiology,* 153**,** 433-443.

RASANATHAN, K., AMERATUNGA, S. & TSE, S. (2006) Asian health in New Zealand - progress and challenges. *New Zealand Medical Journal,* 119.

RAVENSTEIN, E. G. (1885) The laws of migration. *Journal of the Statistical Society,* 48.

RAVENSTEIN, E. G. (1889) The laws of migration. *Journal of the Statistical Society,* 52.

RAYMOND, N. T., JONES, J. R., SWIFT, P. G. F., DAVIES, M. J., LAWRENCE, I. G., MCNALLY, P. G., BURDEN, M. L., GREGORY, R., BOTHA, J. L. & BURDEN, A. C. (2001) Comparative incidence of type I diabetes in children aged under 15 years from South Asian and white or other ethnic backgrounds in Leicestershire, UK, 1989 to 1998. *Diabetologia,* 44**,** B32.

REDAELLI, A., LASKIN, B. L., STEPHENS, J. M., BOTTEMAN, M. F. & PASHOS, C. L. (2005) A systematic literature review of the clinical and epidemiological burden of acute lymphoblastic leukaemia (ALL). *European Journal of Cancer Care,* 14**,** 53-62.

REWERS, M., LAPORTE, R. E., KING, H., TUOMILEHTO, J. & GROUP, D. E. R. I. (1990) Trends in the prevalence and incidence of diabetes: insulin-dependent diabetes mellitus in childhood. *World Health Statistics Quarterly,* 41**,** 179-189.

REWERS, M., NORRIS, J. & DABELEA, D. (2005) Epidemiology of type 1 diabetes. IN EISENBARTH, G. S. (Ed.) *Type 1 diabetes: molecular, cellular, and clinical immunology.* Online Edition Version 2.5 ed., Kluwer Academic.

REYNOLDS, P., ELKIN, E., SCALF, R., VON BEHREN, J. & NEUTRA, R. (2001) A case-control pilot study of traffic exposures and early childhood leukemia using a geographic information system. *Bioelectromagnetics* Suppl 5**,** S58–S68.

REYNOLDS, P., VON BEHREN, J., GUNIER, R., GOLDBERG, D., HERTZ, A. & SMITH, D. (2002) Traffic patterns and childhood incidence rates in California, United States. *Cancer Causes Control,* 13**,** 665–673.

REYNOLDS, P., VON BEHREN, J., GUNIER, R. B., GOLDBERG, D. E., HERTZ, A. & SMITH, D. F. (2003) Childhood cancer incidence rates and hazardous air pollutants in california: an exploratory analysis. *Environmental Health Perspectives,* 111**,** 663-668.

REZAEIAN, M., DUNN, G., ST LEGER, S. & APPLEBY, L. (2006) Ecological association between suicide rates and indices of deprivation in the north west region of England: the importance of the size of the administrative unit. *J Epidemiol Community Health,* 60**,** 956-961.

RHODES, C. J. & ANDERSON, R. M. (1996) Persistence and dynamics in lattice models of epidemic spread. *Journal of Theoretical Biology,* 180**,** 125-133.

RIDOUT, M., HINDE, J. & DEMETRIO, C. G. B. (2001) A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics,* 57**,** 219-223.

ROBERTSON, M. K., RANDLE, M. W. & TUCKER, L. J. (1988) Natural radiation in New Zealand Houses. Wellington, National Radiation Laboratory.

ROBINSON, V. (1996) Introduction: the geographical contribution to the study of human migration. IN ROBINSON, V. (Ed.) *Geography and Migration.* Cheltenham, Edward Elgar Publishing.

ROBSON, B., CORMACK, D. & CRAM, F. (2007) Social and economic indicators. *Hauora: Màori Standards of Health IV. A study of the years 2000–2005.* Wellington, Te Ròpù Rangahau Hauora a Eru Pòmare.

ROCKELL, J. E., GREEN, T. J., SKEAFF, C. M., WHITING, S. J., TAYLOR, R. W., WILLIAMS, S. M., PARNELL, W. R., SCRAGG, R., WILSON, N., SCHAAF, D., FITZGERALD, E. D. & WOHLERS, M. W. (2005) Season and ethnicity are determinants of serum 25-hydroxyvitamin D concentrations in New Zealand children aged 5-14 y. *The Journal of Nutrition,* 135**,** 2602-2608.

ROGERSON, P. A. & HAN, D. (2002) The effects of migration on the detection of geographic differences in disease risk. *Social Science and Medicine,* 55**,** 1817-1828.

ROIVANEN, M., RASILAINEN, S., YLIPAASTO, P., NISSINEN, R., USTINOV, J., BOUWENS, L., EIZIRIK, D. L., HOVI, T. & OTONKOSKI, T. (2000) Mechanisms of coxsackievirus-induced damage to human pancreatic b-cells. *Journal Of Clinical Endocrinology And Metabolism,* 85**,** 432–40.

ROMAN, E., SIMPSON, J., ANSELL, P., KINSEY, S., MITCHELL, C. D., MCKINNEY, P. A., BIRCH, J. M., GREAVES, M., EDEN, T. & ON BEHALF OF THE UNITED KINGDOM CHILDHOOD CANCER STUDY, I. (2007) Childhood acute lymphoblastic leukemia and infections in the first year of life: a report from the United Kingdom Childhood Cancer Study. *American Journal of Epidemiology,* 165**,** 496-504.

ROMMENS, C., LAURIER, D. & SUGIER, A. (2000) Methodology and results of the Nord-Cotentin radioecological study. *Journal of Radiological Protection***,** 361.

ROSENBAUER, J., HERZIG, P., VON KRIES, R., NEU, A. & GIANI, G. (1999) Temporal, seasonal, and geographical incidence patterns of type I diabetes mellitus in children under 5 years of age in Germany. *Diabetologia,* 42**,** 1055-1059.

ROSENBAUM, P. F., BUCK, G. M. & BRECHER, M. L. (2000) Early child-care and preschool experiences and the risk of childhood acute lymphoblastic leukemia. *American Journal of Epidemiology,* 152**,** 1136-1144.

ROSENBERG, M. W. (1998) Medical or health geography? Populations, peoples and places. *International Journal of Population Geography,* 4**,** 211-226.

ROSS, J. A., DAVIES, S. M., POTTER, J. D. & ROBINSON, L. L. (1994) Epidemiology of childhood leukemia, with a focus on infants. *Epidemiology Reviews,* 16**,** 243-272.

ROTHMAN, J. K. & GREENLAND, S. (1998) *Modern epidemiology* Philadelphia Lippincott-Raven.

ROTHWELL, P. M., STAINES, A., SMAIL, P., WADSWORTH, E. & MCKINNEY, P. (1996) Seasonality of birth of patients with childhood diabetes in Britain. *BMJ,* 312**,** 1456-1457.

RUBIN, C. S., HOLMES, A. K., BELSON, M. G., JONES, R. L., FLANDERS, W. D., KIESZAK, S. M., OSTERLOH, J., LUBER, G. E., BLOUNT, B. C., BARR, D. B., STEINBERG, K. K., SATTEN, G. A., MCGEEHIN, M. A. & TODD, R. L. (2007) Investigating Childhood Leukemia in Churchill County, Nevada. *Environmental Health Perspectives,* 115**,** 151-157.

RUDANT, J., BACCAINI, B., RIPERT, M., GOUBIN, A., BELLEC, S., HEMON, D. & CLAVEL, J. (2006) Population-mixing at the place of residence at the time of birth and incidence of childhood leukaemia in France. *European Journal of Cancer,* 42**,** 927-933.

RUDDON, R. W. (2007) *Cancer Biology,* New York, Oxford University Press.

RYTKONEN, M., MOLTCHANOVA, E., RANTA, J., TASKINEN, O., TUOMILEHTO, J. & KARVONEN, M. (2003) The incidence of type 1 diabetes among children in Finland; rural–urban difference. *Health and Place,* 9**,** 315-325.

RYTKONEN, M., RANTA, J., TUOMILEHTO, J. & KARVONEN, M. (2001) Bayesian analysis of geographical variation in the incidence of type I diabetes in Finland. *Diabetologia,* 44**,** B37-B44.

SALMOND, C. & CRAMPTON, P. (2002a) NZDep2001 index of deprivation. Wellington, Department of Public Health, Wellington School of Medicine and Health Sciences.

SALMOND, C. & CRAMPTON, P. (2002b) NZDep2001 index of deprivation user's manual. Wellington, Department of Public Health, Wellington School of Medicine and Health Sciences.

SALMOND, C. E., JOSEPH, J. G., PRIOR, I. A. M., STANLEY, D. G. & WESSEN, A. F. (1985) Longitudinal analysis of the relationship between blood pressure and migration: the Tokelau Island migrant study. *American Journal of Epidemiology,* 122**,** 291-301.

SAMUELSSON, U. & CARSTENSEN, J. (2003) Space-time clustering at birth and at diagnosis of type 1 diabetes mellitus in relation to early clinical manifestation. *Journal of Pediatric Endocrinology & Metabolism,* 16**,** 859-67.

SAMUELSSON, U. & LOFMAN, O. (2004) Geographical mapping of type 1 diabetes in children and adolescents in south east Sweden. *Journal of Epidemiology and Community Health,* 58**,** 388-392.

SAMUELSSON, U. L. F., JOHANSSON, C., CARSTENSEN, J. & LUDVIGSSON, J. (1994) Space-time clustering in insulin-dependent diabetes mellitus (IDDM) in south-east Sweden. *International Journal of Epidemiology,* 23**,** 138-142.

SANTOSA, J. L., CARRASCOB, E., MOOREA, A. L., PÉREZ-BRAVOA, F. & ALBALA, C. (2001) Incidence rate and spatio-temporal clustering of type 1 diabetes in Santiago, Chile, from 1997 to 1998. *Revista de Saúde Pública,* 35 96-100.

SAVILAHTI, E., ÅKERBLOM, H. K., TAINIO, V.-M. & KOSKIMIES, S. (1988) Children with newly diagnosed insulin dependent diabetes mellitus have increased levels of cow's milk antibodies. *Diabetes Research,* 7**,** 137–40.

SCHOBER, E., RAMIA, B., WALDHOER, T. & GROUP, A. D. I. S. (2003) Small area variation in childhood diabetes mellitus in Austria: links to population density, 1989 to 1999. *Journal of Clinical Epidemiology,* 56**,** 269-273

SCHOENLE, E. J., LANG-MURITANO, M., GSCHWEND, S., LAIMBACHER, J., MULLIS, P. E., TORRESANI, T., BIASON-LAUBER, A. & MOLINARI, L. (2001) Epidemiology of type I diabetes mellitus in Switzerland: steep rise in incidence in under 5 year old children in the past decade. *Diabetologia,* 44**,** 286-289.

SCHÜZ, J., GRIGAT, J.-P., BRINKMANN, K. & MICHAELIS, J. (2001a) Residential magnetic fields as a risk factor for childhood acute leukaemia: Results from a German population-based case-control study. *International Journal of Cancer,* 91**,** 728-735.

SCHÜZ, J., GRIGAT, J. P., BRINKMANN, K. & MICHAELIS, J. (2001b) Childhood acute leukaemia and residential 16.7 Hz magnetic fields in Germany. *The British Journal of Cancer,* 84**,** 697.

SCHWIMMBECK, P. L., DYRBERG, T. & OLDSTONE, M. B. (1988) Abrogation of diabetes in BB rats by acute virus infection. Association of viral-lymphocyte interactions. *The Journal of Immunology,* 140**,** 3394-3400.

SCOTT, F. W., NORRIS, J. M. & KOLB, H. (1996) Milk and type I diabetes: examining the evidence and broadening the focus. *Diabetes Care,* 19**,** 379.

SCOTT, R. S., BROWN, L. J., DARLOW, B. A., FORBES, L. V. & MOORE, M. P. (1992) Temporal variation in incidence of IDDM in Canterbury, New Zealand. *Diabetes Care,* 15**,** 895-899.

SHAH, A. & COLEMAN, M. P. (2007) Increasing incidence of childhood leukaemia: a controversy re-examined. *British Journal of Cancer,* 97**,** 1009-1012.

SHAVER, K. A. & AL, E. (1985) Congenital rubella syndrome and diabetes: a review of epidemiologic, genetic, and immunological factors. *American Annals Of The Deaf,* 130**,** 526-532.

SHAW, G. & WHEELER, D. (1994) *Statistical techniques in geographical analysis.,* London, David Fulton.

SHOULS, S., CONGDON, P. & CURTIS, S. (1996) Modelling inequality in reported long term illness in the UK: combining individual and area characteristics. *J Epidemiol Community Health,* 50**,** 366-376.

SHRAPNEL, W. & TRUSWELL, S. (2006) Vitamin D deficiency in Australia and New Zealand: What are the dietary options? *Nutrition & Dietetics,* 63**,** 206-212.

SHU, X. O., HAN, D., SEVERSON, R. K., CHEN, Z. & ET AL. (2002) Birth characteristics, maternal reproductive history, hormone use during pregnancy, and risk of childhood acute lymphocytic leukemia by immunophenotype (United States). *Cancer Causes & Control,* 13**,** 15.

SHU, X. O., LINET, M. S., STEINBUCH, M., WEN, W. Q., BUCKLEY, J. D., NEGLIA, J. P., POTTER, J. D., REAMAN, G. H. & ROBISON, L. L. (1999) Breast-Feeding and risk of childhood acute leukemia. *Journal of the National Cancer Institute,* 91**,** 1765-1772.

SHURTLEFF, S. A., BUIJS, A., BEHM, F. G., RUBNITZ, J. E., RAIMONDI, S. C., HANCOCK, M. L., CHAN, G. C., PUI, C. H., GROSVELD, G., DOWNING, J. R. & AL., E. (1995) TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. *Leukemia,* 9**,** 1985–1989.

SIEMIATYCKI, J., COLLE, E., CAMPBELL, S., DEWAR, R. A. & BELMONTE, M. M. (1989) Case-control study of IDDM. *Diabetes Care,* 12**,** 209-216.

SILVEY, R. (2004a) On the boundaries of a subfield: social theory's incorporation into population geography. *Population, Space and Place,* 10**,** 303-308.

SILVEY, R. (2004b) Power, difference and mobility: feminist advances in migration studies. *Progress in Human Geography,* 28**,** 490-506.

SIMPSON, N. E. (1976) A review of family data. IN CREUTZFELDT, W., KOBBERLING, J. & NEEL, J. V. (Eds.) *The genetics of diabetes mellitus* Berlin, Springer.

SINGH, B., PRANGE, S. & JEVNIKAR, A. M. (1998) Protective and destructive effects of microbial infection in insulin-dependent diabetes mellitus. *Seminars in Immunology,* 10**,** 79-86.

SINGH, G., K. & MILLER, B., A. (2004) Health, life expectancy, and mortality patterns among immigrant populations in the United States. *Canadian Journal of Public Health,* 95**,** I14.

SINKS, T., SMITH, M., CARROLL, W. L., HASTINGS, C., RIVERS, C., LINDEMAN, R. & DAHL-CURTIS, R. (2004) Final report and recommendations to the Nevada State Health Division.

SKELLY, C., BLACK, W., HEARNDEN, M., EYLES, R. & WEINSTEIN, P. (2002) Disease surveillance in rural communities is compromised by address geocoding uncertainty: a case study of campylobacteriosis. *Australian Journal of Rural Health,* 10**,** 87-93.

SLAMA, G. (2003) Type 1 diabetes: an overview. IN PICKUP, J. C. & WILLIAMS, G. (Eds.) *Textbook of diabetes 1.* 3rd ed. Oxford, Blackwell Publishing.

SMALLMAN-RAYNOR, M. & CLIFF, A. D. (1999) The spatial dynamics of epidemic diseases in war and peace: Cuba and the insurrection against Spain, 1895-98. *Transactions of the Institute of British Geographers,* 24**,** 331-352.

SMALLMAN-RAYNOR, M. R., MUIR, K. R. & SMITH, S. J. (1998) The geographical assignment of cancer units: patient accessibility as an optimal allocation problem. *Public Health,* 112**,** 379-383.

SMITH, A., ROMAN, E. & SIMPSON, J. (2007) Childhood leukaemia and socioeconomic status. *International Journal of Epidemiology,* 36**,** 1156.

SMITH, A., ROMAN, E., SIMPSON, J., ANSELL, P., FEAR, N. T. & EDEN, T. (2006) Childhood leukaemia and socioeconomic status: fact or artefact? A report from the United Kingdom childhood cancer study (UKCCS). *Int. J. Epidemiol.***,** dyl193.

SMITH, D. G., BARTLEY, M. & BLANE, D. (1990) The Black report on socioeconomic inequalities in health 10 years on. *British Medical Journal,* 301**,** 373–377.

SMITH, G. D. & PHILLIPS, A. N. (1992) Confounding in epidemiological studies: why "independent" effects may not be all they seem. *British Medical Journal,* 305**,** 757–759.

SMITH, R. B. (1987) Diabetes in Young New Zealanders: results of national survey 1978-82. *New Zealand Medical Journal,* 100**,** 581-584.

SMULEVICH, V. B., SOLIONOVA, L. G. & BELYAKOVA, S. V. (1999) Parental occupation and other factors and cancer risk in children: I. Study methodology and non-occupational factors. *International Journal of Cancer,* 83**,** 712-717.

SMYTH, F. & THOMAS, R. (2005) `Burning Issues': an introduction to selected papers from the 10th International Symposium in Medical Geography, Manchester 2003. *Social Science & Medicine,* 60**,** 2657-2660.

SONGINI, M., CASU, A., ASHKENAZI, I. & LARON, Z. (2001) Seasonality of birth in children (0-14 years) and young adults (0-29 years) with type 1 diabetes mellitus in Sardinia differs from that in the general population. *Journal of Pediatric Endocrinology and Metabolism,* 14**,** 781-3.

STAINES, A. (1996) The geographical epidemiology of childhood insulin independent diabetes and childhood acute lymphoblastic leukaemia in Yorkshire. University of Leeds.

STAINES, A. (2001) Commentary: population mixing and childhood diabetes. *International Journal of Epidemiology,* 30**,** 538-539.

STAINES, A., BODANSKY, H. J., MCKINNEY, P. A., ALEXANDER, F. E., MCNALLY, R. J. Q., LAW, G. R., LILLEY, H. E. B., STEPHENSON, C. & CARTWRIGHT, R. (1997) Small area variation in the incidence of childhood insulin-dependent diabetes mellitus in Yorkshire, UK: links with overcrowding and population density. *International Journal of Epidemiology,* 26**,** 1307-1313.

STATISTICS NEW ZEALAND (1997) 1996 Census of population and dwellings: an introduction to the census. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2001) 2001 Census of population and dwellings: an introduction to the census. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2004) New Zealand Family and Household Projections: 2001(base)–2021. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2006a) New Zealand: an urban/rural profile. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2006b) QuickStats about population mobility: 2006 census. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2006c) QuickStats national highlights: 2006 census. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2007a) Demographic trends (2006) - reference report. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2007b) Geographic hierarchy. Wellington, Statistics New Zealand.

STATISTICS NEW ZEALAND (2007c) Tourism and Migration 2006. Wellington, Statistics New Zealand.

STEFFEN, C., AUCLERC, M. F., AUVRIGNON, A., LEVERGER, G., VILMER, E., HEMON, D. & CLAVEL, J. (2004) Acute childhood leukaemia and environmental exposure to potential sources of benzene and other hydrocarbons; a case-control study. *Occupational and Environmental Medicine,* 61**,** 773(6).

STEINMAUS, C., LU, M., TODD, R. L. & SMITH, A. H. (2004) Probability estimates for the unique childhood leukemia cluster in Fallon, Nevada, and risks near other U.S. military aviation facilities *Environmental Health Perspectives,* Vol.112**,** 766-771.

STELIAROVA-FOUCHER, E., STILLER, C., KAATSCH, P., BERRINO, F. & COEBERGH, J.-W. (2005) Trends in childhood cancer incidence in Europe, 1970-99. *The Lancet,* 365**,** 2088.

STELIAROVA-FOUCHER, E., STILLER, C., KAATSCH, P., BERRINO, F., COEBERGH, J.-W., LACOUR, B. & PERKIN, M. (2004) Geographical patterns and time trends of cancer incidence and survival among children and adolescents in Europe since the 1970s (the ACCIS project): an epidemiological study. *The Lancet,* 364**,** 2097-2105.

STERNE, J. A. C., SMITH, G. D. & COX, D. R. (2001) Sifting the evidence - what's wrong with significance tests? Another comment on the role of statistical methods. *British Medical Journal,* 322**,** 226-231.

STEWART, A., WEBB, J. & HEWITT, D. (1958) A Survey of Childhood Malignancies. *British Medical Journal,* 1**,** 1495–1508.

STILLER, C. A. (2004) Epidemiology and genetics of childhood cancer. *Oncogene,* 23**,** 6429-6444.

STILLER, C. A. & BOYLE, P. J. (1996) Effect of population mixing and socioeconomic status in England and Wales, 1979-85, on lymphoblastic leukaemia in children. *British Medical Journal,* 313**,** 1297-1300.

STILLER, C. A. & PARKIN, D. M. (1996) Geographic and ethnic variations in the incidence of childhood cancer. *British Medical Bulletin,* 52**,** 682-703.

STRACHAN, D. P. (1989) Hay fever, hygiene, and household size. *British Medical Journal,* 299**,** 1259–1260.

SYME, S. L., MARMOT, M. G., KAGAN, A., KATO, H. & RHOADS, G. (1975) Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: introduction. *American Journal of Epidemiology, ,* 102**,** 477-480. .

TANGO, T. & TAKAHASHI, K. (2005) A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics* 4**,** 11.

TAPLIN, C. E., CRAIG, M. E., LLOYD, M., TAYLOR, C., CROCK, P., SILINK, M. & HOWARD, N. J. (2005) The rising incidence of childhood type 1 diabetes in New South Wales, 1990–2002. *The Medical Journal of Australia,* 183**,** 243-246.

TEDESCHI, A. & AIRAGHI, L. (2006) Is affluence a risk factor for bronchial asthma and type 1 diabetes? *Pediatric Allergy and Immunology,* 17**,** 533-537.

TERRACINI, B. & MAULE, M. M. (2007) Aetiological clues from the descriptive epidemiology of childhood acute lymphatic leukaemia and other malignancies. *Journal of Epidemiology and Community Health,* 61**,** 180-181.

THUN, M. J. & SINKS, T. (2004) Understanding cancer clusters. *CA: A Cancer Journal for Clinicians,* 54**,** 273-280.

TILLING, K., PETERS, T. & STERNE, J. (2005) Key issues in the statistical analysis of quantitative data in research on health and health services. IN BOWLING, A. & EBRAHIM, S. (Eds.) *Handbook of Health Research Methods: Investigation, Measurement, and Analysis.* Berkshire, Open University Press.

TROIANO, R. P., FLEGAL, K. M., KUCZMARSKI, R. J., CAMPBELL, S. M. & JOHNSON, C. L. (1995) Overweight prevalence and trends for children and adolescents. The national health and nutrition examination surveys, 1963 to 1991. *Archives of Pediatrics & Adolescent Medicine,* 149**,** 1085-1091.

TRUSWELL, A. S. (2005) The A2 milk case: a critical review. *European Journal of Clinical Nutrition,* 59**,** 623-631

TUKUITONGA, C. F., SOLOMON, N. & STEWART, A. (1992) Incidence of cancer among Pacific Island people in New-Zealand. *New Zealand Medical Journal,* 105**,** 463-466.

TUNSTALL, H. V. Z., SHAW, M. & DORLING, D. (2004) Places and health *Journal of Epidemiology & Community Health,* 58**,** 6-10.

TUOMILEHTO, J., KARVONEN, M., PITKÄNIEMI, J., VIRTALA, E., KOHTAMÄKI, K., TOIVANEN, L. & TUOMILEHTO-WOLF, E. (1999) Record-high incidence of Type I (insulin-dependent) diabetes mellitus in Finnish children. *Diabetologia,* 42**,** 655-660.

TUOMILEHTO, J., VIRTALA, E., KARVONEN, M., LOUNAMAA, R., PITKÄNIEMI, J., REUNANEN, A., TUOMILEHTO-WOLF, E., TOIVANEN, L. & GROUP, D. S. (1995) Increase in incidence of insulin-dependent diabetes mellitus among children in Finland. *International Journal of Epidemiology,* 24**,** 984-992.

TURNBULL, A., BARRY, D., WICKENS, K. & CRANE, J. (2004) Changes in body mass index in 1112-year-old children in Hawkes Bay, New Zealand (1989-2000). *Journal of Paediatrics and Child Health,* 40**,** 33-37.

UKCCS INVESTIGATORS (2000) Childhood cancer and residential proximity to power lines. *The British Journal of Cancer,* 83**,** 1573.

UKCCS INVESTIGATORS (2001) Breastfeeding and childhood cancer. *British Journal of Cancer,* 85**,** 1685-1694.

UKCCS INVESTIGATORS (2002a) The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: 1: radon gas. *The British Journal of Cancer,* 86**,** 1721.

UKCCS INVESTIGATORS (2002b) The United Kingdom Childhood Cancer Study of exposure to domestic sources of ionising radiation: 2: gamma radiation. *The British Journal of Cancer,* 86**,** 1727.

UKCCS INVESTIGATORS, SKINNER, J., MEE, T. J., BLACKWELL, R. P., MASLANYJ, M. P., SIMPSON, J., ALLEN, S. G. & DAY, N. E. (2002) Exposure to power frequency electric fields and the risk of childhood cancer in the UK. *The British Journal of Cancer,* 87**,** 1257.

URQUHART, J. D., BLACK, R. J., MUIRHEAD, M. J., SHARP, L., MAXWELL, M., EDEN, O. B. & JONES, D. A. (1991) Case-control study of leukaemia and non-Hodgkin's

lymphoma in children in Caithness near the Dounreay nuclear installation. *British Medical Journal,* 302**,** 687-92.

VAN DER WERF, N., KROESE, F. G. M., ROZING, J. & HILLEBRANDS, J. L. (2007) Viral infections as potential triggers of type 1 diabetes. *Diabetes/Metabolism Research and Reviews,* 23**,** 169-183.

VAN LENTHE, F. J., MARTIKAINEN, P. & MACKENBACH, J. P. (2007) Neighbourhood inequalities in health and health-related behaviour: results of selective migration? *Health & Place,* 13**,** 123-137.

VAN MAANEN, J. M. S., ALBERING, H. J., VAN BREDA, S. G. J. & *AL.*, *E.* (1999) Nitrate in drinking water and risk of childhood diabetes in the Netherlands [Letter]. *Diabetes Care,* 22**,** 1750.

VERGE, C. F., HOWARD, N. J., IRWIG, L., SIMPSON, J. M., MACKERRAS, D. & SILINK, M. (1994) Environmental factors in childhood IDDM. A population-based, case-control study. *Diabetes Care,* 17**,** 1381-1389.

VIEL, J. F., POBEL, D. & CARRÉ, A. (1995) Incidence of leukaemia in young people around the La hague nuclear waste reprocessing plant: A sensitivity analysis. *Statistics in Medicine,* 14**,** 2459-2472.

VIRTANEN, S. M., JAAKKOLA, L., RÄSÄNEN, L., YLÖNEN, K., ARO, A., LOUNAMAA, R., AKERBLOM, H. K. & TUOMILEHTO, J. (1994a) Nitrate and nitrite intake and the risk for type 1 diabetes in Finnish children. Childhood Diabetes in Finland Study Group. *Diabetic Medicine,* 11**,** 656-62.

VIRTANEN, S. M. & KNIP, M. (2003) Nutritional risk predictors of {beta} cell autoimmunity and type 1 diabetes at a young age. *Am J Clin Nutr,* 78**,** 1053-1067.

VIRTANEN, S. M., RÄSÄNEN, L., ARO, A., LINDSTROM, J., SIPPOLA, H., LOUNAMAA, R., TOIVANEN, L., TUOMILEHTO, J. & AKERBLOM, H. K. (1991) Infant feeding in Finnish children < 7 yr of age with newly diagnosed IDDM. *Diabetes Care,* 14**,** 415-417.

VIRTANEN, S. M., SAUKKONEN, T., SAVILAHTI, E. & ET AL. (1994b) Diet, cow's milk protein antibodies and the risk of IDDM in Finnish children. *Diabetologia,* 37**,** 381-387.

VISKARI, H., LUDVIGSSON, J., UIBO, R., SALUR, L., MARCIULIONYTE, D., HERMANN, R., SOLTESZ, G., FÜCHTENBUSCH, M., ZIEGLER, A. G., KONDRASHOVA, A., ROMANOV, A., KAPLAN, B., LARON, Z., KOSKELA, P., VESIKARI, T., HUHTALA, H., KNIP, M. & HYÖTY, H. (2005) Relationship between the incidence of type 1 diabetes and maternal enterovirus antibodies: time trends and geographical variation. *Diabetologia,* 48**,** 1280.

VUONG, Q. H. (1989) Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica,* 57**,** 307-333.

WADSWORTH, E. J. K., SHIELD, J. P. H., HUNT, L. P. & BAUM, J. D. (1997) A case-control study of environmental factors associated with diabetes in the under 5s. *Diabetic Medicine,* 14**,** 390-396.

WAGENKNECHT, L. E., ROSEMAN, J. M. & HERMAN, W. H. (1991) Increased incidence of insulin-dependent diabetes mellitus following an epidemic of coxsackievirus B5. *American Journal of Epidemiology,* 133**,** 1024-1031.

WAKEFORD, R. (2004) The cancer epidemiology of radiation. *Oncogene,* 23**,** 6404-6428.

WALTON-ROBERTS, M. (2004) Transnational migration theory in population geography: gendered practices in networks linking Canada and India. *Population, Space and Place,* 10**,** 361-373.

WARD, C. & MASGORET, A. M. (2008) Attitudes toward Immigrants, immigration, and multiculturalism in New Zealand: a social psychological analysis. *International Migration Review,* 42**,** 227-248.

WARD, G. (1917) The infective theory of acute leukaemia. *British Journal of Children's Diseases,* 14**,** 10.

WARD, R. H., CHIN, P. G. & PRIOR, I. (1980) Tokelau Island migrant study, effect of migration on the familial aggregation of blood pressure. *Hypertension,* 2**,** 43-54.

WARFA, N., BHUI, K., CRAIG, T., CURTIS, S., MOHAMUD, S., STANSFELD, S., MCCRONE, P. & THORNICROFT, G. (2006) Post-migration geographical mobility, mental health and health service utilisation among Somali refugees in the UK: A qualitative study. *Health & Place,* 12**,** 503-515.

WARTENBERG, D., SCHNEIDER, D. & BROWN, S. (2004) Childhood leukaemia incidence and the population mixing hypothesis in US SEER data. *British Journal of Cancer,* 90**,** 1771-1776.

WAUGH, N. R. (1986) Insulin-dependent diabetes in a Scottish region: incidence and urban/rural differences. *Journal of Epidemiology and Community Health,* 40**,** 240-243.

WEBB, P., BAIN, C. & PIROZZO, S. (2005) *Essential epidemiology: an introduction for students and health professionals,* New York, Cambridge University Press.

WEISS, R. A. & MCMICHAEL, A. J. (2004) Social and environmental risk factors in the emergence of infectious diseases. *Nature Medicine,* 10.

WELTON, T. A. (1872) The effects of migration in disturbing local rates of mortality as exemplifed in the statistics of London and the surrounding country for the years 1851-60. *Journal of the Institute of Actuaries,* 16**,** 153-186.

WHITE, P. & JACKSON, P. (1995) (Re)theorising population geography. *International Journal of Population Geography,* 1**,** 111-23.

WHITE, S. (1980) A philosophical dichotomy in migration research. *Professional Geographer,* 32**,** 6-13.

WHO (2006) SARS how a global epidemic was stopped. Manila, WHO Western Pacific Region.

WIEMELS, J. L., CAZZANIGA, G., DANIOTTI, M., EDEN, O. B., ADDISON, G. M., MASERA, G., SAHA, V., BIONDI, A. & GREAVES, M. F. (1999) Prenatal origin of acute lymphoblastic leukaemia in children. *The Lancet,* 354**,** 1499-1503.

WILKIN, T. J. (2001) The accelerator hypothesis: weight gain as the missing link between type I and type II diabetes. *Diabetologia,* 44**,** 914-922.

WILLIS, J. A., SCOTT, R. S. & DARLOW, B. A. (2005) To: Campbell-Stokes, PL, Taylor BJ on behalf of the New Zealand Children's Diabetes Working Group (2005). Prospective incidence study of diabetes mellitus in New Zealand children aged 0 to 14 years. Diabetologia 48: 643-648. *Diabetologia,* 48**,** 2442-2443.

WILLIS, J. A., SCOTT, R. S., DARLOW, B. A., LEWY, H., ASHKENAZI, I. & LARON, Z. (2002a) Seasonality of birth and onset of clinical disease in children and adolescents (0-19 years) with type 1 diabetes mellitus in Canterbury, New Zealand. *Journal of Pediatric Endocrinology & Metabolism,* 15**,** 645-647.

WILLIS, J. A., SCOTT, R. S., DARLOW, B. A., LUNT, H. & MOORE, E. (1997) Hepatitus B immunisation and the epidemiology of IDDM in children and adolescents aged <20 years in Canterbury, NZ. *Diabetes.*

WILLIS, J. A., SCOTT, R. S., DARLOW, B. A., NESBIT, J. W., ANDERSON, P., MOORE, M. P., LUNT, H. & COLE, D. R. (2002b) Incidence of type 1 diabetes mellitus diagnosed before age 20 years in Canterbury, New Zealand over the last 30 years. *Journal of Pediatric Endocrinology & Metabolism,* 15**,** 637-643.

WILLIS, J. A., SCOTT, R. S., DARLOW, D. A., ELLIOT, R. B. & MCGREGOR, M. (2004) Type 1 diabetes in children and adolescents: Auckland and Canterbury Longitudinal Cohorts. *New Zealand Society for the Study of Diabetes 28th Annual Meeting.* Auckland.

WILSON, G., MILES, A. & PARKER, M. T. (Eds.) (1983) *Principles of bacteriology, virology and immunity, vol 1,* London, Edward Arnold.

WILSON, I. V., GALE, E. A. M. & BINGLEY, P. J. (2007) The rising incidence of childhood Type 1 diabetes in the Oxford region 1985-2004. *Diabetic Medicine,* 24**,** 24-24.

WILTON, R. D. (1999) Qualitative health research: negotiating life with HIV/AIDS. *The Professional Geographer,* 51**,** 254-264.

WONG, D. W. S. (2003) Spatial decomposition of segregation indices: a framework toward measuring segregation at multiple levels. *Geographical Analysis,* 35 179-195.

WOOD, G. R. (2002) Generalised linear accident models and goodness of fit testing. *Accident Analysis and Prevention* 34**,** 417–427.

WOODFORD, K. B. (2006) Letter to the editor: A critique of Truswell's A2 milk review. *European Journal of Clinical Nutrition,* 60**,** 437-439.

WORLD HEALTH ORGANIZATION (2007) Electromagnetic fields and public health: exposure to extremely low frequency fields. *Fact sheet number 322.* Geneva, World Health Organisation.

WORLD HEALTH ORGANIZATION & DIAMOND PROJECT GROUP ON EPIDEMICS (1992) Childhood diabetes, epidemics, and epidemiology: an approach for controlling diabetes. *American Journal of Epidemiology,* 135**,** 803-816.

WU, D., KENDALL, D., LUNT, H., WILLIS, J. A., DARLOW, B. A. & FRAMPTON, C. (2005) Prevalence of type 1 diabetes in New Zealanders aged 0-24 years. *The New Zealand Medical Journal,* 118**,** 1557-1563.

XIE, Y., DAVIES, S. M., XIANG, Y., ROBISON, L. L. & ROSS, J. A. (2003) Trends in leukemia incidence and survival in the United States (1973-1998). *Cancer,* 97**,** 2229-2235.

YOON, J. W. & JUN, H. S. (2003) Viruses in the pathogenesis of type 1 diabetes. IN PICKUP, J. C. & WILLIAMS, G. (Eds.) *Textbook of diabetes 1*. 3rd ed. Oxford, Blackwell Publishing.

YOON, J. W., MCCLINTOCK, P. R., BACHURSKI, C. J. & AL:, E. (1985) Virus induced diabetes mellitus. No evidence for immune mechanisms in the destruction of b-cells by the D-variant of encephalomyocarditis virus. *Diabetes,* 34**,** 922–25.

YOSHIMOTO, Y., NEEL, J. V., SCHULL, W. J., KATO, H., SODA, M., ETO, R. & MABUCHI, K. (1990) Malignant tumors during the first 2 decades of life in the offspring of atomic bomb survivors. *American journal of human genetics,* 46**,** 1041–1052.

ZELINSKY, W. (1971) Hypothesis of mobility transition. *Geographical Review,* 61**,** 219-249.

ZIEGLER, A.-G., SCHMID, S., HUBER, D., HUMMEL, M. & BONIFACIO, E. (2003) Early infant feeding and risk of developing type 1 diabetes-associated autoantibodies. *The Journal of the American Medical Association,* 290**,** 1721-1728.

ZIEGLER, A. G., HUMMEL, M., SCHENKER, M. & BONIFACIO, E. (1999) Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. *Diabetes,* 48**,** 460-468.

ZLOTNIK, H. (1999) Trends of international migration since 1965: what existing data reveal. *International Migration,* 37**,** 21-61.